Closed-form Output Statistics of MIMO Block-Fading Channels

Giuseppa Alfano, *Member, IEEE,* Carla-Fabiana Chiasserini, *Senior Member, IEEE,* Alessandro Nordio, *Member, IEEE,* Siyuan Zhou

***Abstract*—The information that can be transmitted through a wireless channel, with multiple-antenna equipped transmitter and receiver, is crucially influenced by the channel behavior as well as by the structure of the input signal. We characterize in closed form the probability density function (pdf) of the output of MIMO block-fading channels, for an arbitrary SNR value. Our results provide compact expressions for such output statistics, paving the way to a more detailed analytical information-theoretic exploration of communications in presence of block fading. The analysis is carried out assuming two different structures for the input signal: the i.i.d. Gaussian distribution and a product form that has been proved to be optimal for non-coherent communication, i.e., in absence of any channel state information. When the channel is fed by an i.i.d. Gaussian input, we assume the Gramian of the channel matrix to be unitarily invariant and derive the output statistics in both the noise-limited and the interference-limited scenario, considering different fading distributions. When the product-form input is adopted, we provide the expressions of the output pdf as the relationship between the overall number of antennas and the fading coherence length varies. We also highlight the relation between our newly derived expressions and the results already available in the literature, and, for some cases, we numerically compute the mutual information, based on the proposed expression of the output statistics.**

Index terms: Output statistics, MIMO, block fading, imperfect channel state information.

I. Introduction

The availability of an explicit statistical characterization of the output of a wireless channel, impaired by additive and multiplicative random disturbance, is of paramount importance to communication- and information-theoretic purposes. Indeed, a closed-form expression for the output probability density function (pdf) is relevant for the evaluation of the ergodic mutual information between the input and the output signals of a randomly faded channel [1]. It also turns out to be crucial in the finite block-length regime, in order to characterize the information density of the communication at hand [2].

In spite of its importance, few explicit results are available in the literature for the output signal pdf in the case of MIMO block-independent fading channels. The works in [3], [4], [5] all focus on the case of block-Rayleigh fading. In these papers, the output statistics are derived under different assumptions on the relative values of the number of involved antennas and of the coherence length of the fading. The input distribution, too, plays a crucial role in the cited derivations. More specifically, in [3] the authors assume the input to be i.i.d. Gaussian and investigate the behavior of the output distribution as the fading coherence length varies from being quite short to very long, compared to the overall number of transmit and receive antennas. In both [4] and [5], instead, the input is assumed to be given by the product of a diagonal matrix (representing the power allocation over the transmit antennas) times an isotropically distributed matrix with unitary columns. The main difference between the two papers is in the assumption on the fading duration. Indeed, the first one focuses on the case where the coherence length of the Rayleigh fading is greater than the number of involved antennas; in this case, the high Signal to Noise Ratio (SNR)-optimal power allocation matrix turns out to be a scaled version of the identity matrix [6]. The study in [5], instead, solves the problem of characterizing, again in the high-SNR regime, the optimal power allocation profile, assuming the fading coherence length to be shorter, compared to the number of involved antennas. In the latter case, indeed, the diagonal matrix of the power allocation is characterized by the eigenvalues of a matrix-variate Beta joint distribution of the entries [5].

In this paper, we consider both the input models described above, and derive closed form expressions of the output pdf in presence of a multiple-antenna channel affected by additive noise and block-fading. In particular, in the case of i.i.d. Gaussian input, our procedure allows the derivation of a closed-form expression for the output statistics of channels with unitarily invariant fading law. Apart from the canonical i.i.d. Rayleigh fading, already treated in [3], this encompasses the Rician channel with scalar Line-of-Sight (LOS) matrix, whose analysis was previously limited to the evaluation of the fading number [7], and the LOS MIMO [8] with a certain amount of residual scattering. Also, we provide results for the Land Mobile Satellite (LMS) with scalar average power LOS matrix [9, Property I] and for the above cases of MIMO Rayleigh and Rician fading communications impaired by Rayleigh-faded co-channel interference [10]. We remark that the expressions of the output pdf that we derive hold for any arbitrary value of SNR.

The paper is organized as follows. Section II introduces the notations used throughout the paper, the communication model of the wireless system and relevant mathematical background. Section III presents the analytical derivation of the output pdfs, as the channel is fed by an i.i.d. Gaussian input. Section IV provides the output pdfs in presence of optimized product

Alessandro Nordio is with IEIIT-CNR (Institute of Electronics, Telecommunications and Information Engineering of the National Research Council of Italy), Italy, email: alessandro.nordio@ieiit.cnr.it.

Giuseppa Alfano, Siyuan Zhou, and C.-F. Chiasserini are with the Dipartimento di Elettronica and Telecomunicazioni, Politecnico di Torino, Torino, Italy, email: alfano@tlc.polito.it, siyuan.zhou@polito.it, chiasserini@polito.it.

form in Rayleigh block-fading channels. Finally, Section V concludes the paper.

II. Preliminaries and Communication Model

A. Notations

1) Vectors and matrices: Throughout the paper, uppercase and lowercase boldface letters denote matrices and vectors, respectively. The identity matrix is denoted by $\mathbf{I}$. The pdf of a random matrix $\mathbf{A}$, $p_{\mathbf{A}}(\mathbf{A})$, is simply indicated with $p(\mathbf{A})$, except when referring to $\mathbf{A}$ is needed for clarity. $\mathbb{E}[\cdot]$ represents statistical expectation, $(\cdot)^{\mathsf{H}}$ indicates the conjugate transpose operator, $\mathsf{Tr}\{\cdot\}$ denotes the trace of a square matrix, and $\|\cdot\|$ stands for the Euclidean norm[1]. Also, we indicate with $\{a_{ij}\}$ the matrix whose elements are a_{ij} and with $|\mathbf{A}|$, or $|\{a_{ij}\}|$, the determinant of matrix $\mathbf{A}$. We often employ the following property of the determinant:

Property 1: Let $\mathbf{F} = \{f_{ij}\}$ be an $m \times m$ matrix where $f_{ij} = \alpha a_i b_{ij} c_j$. Then,

$$|\mathbf{F}| = \alpha^m |\{b_{ij}\}| \prod_{i=1}^{m} a_i \prod_{j=1}^{m} c_j \,. \tag{1}$$

2) Complex multivariate Gamma function: $\Gamma_m(a)$ is the complex multivariate Gamma function defined as [11]:

$$\Gamma_m(a) = \pi_m \prod_{\ell=1}^{m} \Gamma(a-\ell+1)$$

with m being a non-negative integer and

$$\pi_m = \pi^{m(m-1)/2} \,.$$

3) Vandermonde determinant: Let $\mathbf{A}$ be an $m \times m$ Hermitian matrix with eigenvalues $a_1, \ldots, a_m$. Then the Vandermonde determinant of $\mathbf{A}$ is defined as [12, eq. (2.10)]:

$$\mathcal{V}(\mathbf{A}) = \prod_{1 \le i < j \le m} (a_i - a_j) \,, \tag{2}$$

where we assume the eigenvalues to be ordered in decreasing order so that $\mathcal{V}(\mathbf{A})$ is non negative. Moreover, for any constant c, we have $\mathcal{V}(c\mathbf{A}) = c^{m(m-1)/2}\mathcal{V}(\mathbf{A})$. Let also $\mathbf{F} = \{f_i(a_j)\}$, $i, j = 1, \ldots, m$, be an $m \times m$ matrix, where $f_i(\cdot)$'s are any differentiable functions. Clearly, if the eigenvalues of $\mathbf{A}$ are not distinct, $\mathcal{V}(\mathbf{A}) = 0$ and $|\mathbf{F}| = 0$. In such a case, the ratio $|\mathbf{F}|/\mathcal{V}(\mathbf{A})$, which appears in the density of many matrices that we study in the following, can be evaluated by applying l'Hôpital's rule. More precisely, let n be an integer such that $0 < n < m$, then [13, Lemma 5]

$$\lim_{a_{n+1}, \ldots, a_m \to a} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{A})} = \frac{\pi_m \Gamma_n(m)}{\pi_n \Gamma_m(m)} \frac{|\widetilde{\mathbf{F}}|}{\mathcal{V}(\widetilde{\mathbf{A}})} |\widetilde{\mathbf{A}} - a\mathbf{I}|^{n-m} \tag{3}$$

where $\widetilde{\mathbf{A}}$ is of size $n \times n$ and has eigenvalues $a_1, \ldots, a_n$ and

$$(\widetilde{\mathbf{F}})_{ij} = \begin{cases} f_i(a_j) & i = 1, \ldots, m; \; j = 1, \ldots, n \\ f_i^{(m-j)}(a) & i = 1, \ldots, m; \; j = n+1, \ldots, m \end{cases}$$

[1]As applied to a matrix, we mean $\|\mathbf{A}\|^2 = \mathsf{Tr}\{\mathbf{A}^{\mathsf{H}}\mathbf{A}\}$.

with $f_i^{(k)}(\cdot)$ denoting the k-th derivative of $f_i(\cdot)$. For $n = 0$, we have

$$\lim_{a_1, \ldots, a_m \to a} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{A})} = \frac{\pi_m}{\Gamma_m(m)} |\widetilde{\mathbf{F}}| \tag{4}$$

where $(\widetilde{\mathbf{F}})_{ij} = f_i^{(m-j)}(a)$, $i, j = 1, \ldots, m$.

4) Generalized hypergeometric function: The generalized hypergeometric function is defined as ${}_pF_q(\mathbf{a}; \mathbf{b}; \mathcal{X})$, where $\mathbf{a} = [a_1, \ldots, a_p]^{\mathsf{T}}$, $\mathbf{b} = [b_1, \ldots, b_q]^{\mathsf{T}}$, and $\mathcal{X}$ is a set of arguments that can be either scalars or square matrices [14]. In the case of a single scalar argument, $\mathcal{X} = \{x\}$, the generalized hypergeometric function is defined as in [12, eq. (2.24)]:

$${}_pF_q(\mathbf{a}; \mathbf{b}; x) = \sum_{k=0}^{\infty} \frac{[\mathbf{a}]_k}{[\mathbf{b}]_k} \frac{x^k}{k!} \tag{5}$$

where $[\mathbf{a}]_k = \prod_{i=1}^{p} [a_i]_k$, $[\mathbf{b}]_k = \prod_{j=1}^{q} [b_j]_k$, and $[z]_k = \Gamma(z+k)/\Gamma(z)$ denotes the Pochhammer symbol. Note that ${}_0F_0(;;x) = \mathrm{e}^x$, and ${}_1F_0(a;;x) = (1-x)^{-a}$. The function ${}_0F_1(;b;x)$ is closely related to the Bessel's function, and in the literature functions ${}_1F_1(a;b;x)$ and ${}_2F_1(a_1, a_2; b; x)$ are also called *confluent hypergeometric function of the first kind* and *Gauss's hypergeometric function*, respectively.

The generalized hypergeometric function of two matrix arguments, $\mathcal{X} = \{\mathbf{\Phi}, \mathbf{\Psi}\}$, both of size $m \times m$, can be written through hypergeometric functions of scalar arguments as [12, eq. (2.34)]

$${}_pF_q(\mathbf{a}; \mathbf{b}; \mathbf{\Phi}, \mathbf{\Psi}) = c \frac{|\{{}_pF_q(\tilde{\mathbf{a}}; \tilde{\mathbf{b}}; \phi_h \psi_k)\}|}{\mathcal{V}(\mathbf{\Phi})\mathcal{V}(\mathbf{\Psi})} \tag{6}$$

$h, k = 1, \ldots, m$, where the constant c is given by [12]

$$c = \frac{\Gamma_m(m)}{\pi_m^{q-p+1}} \left[\prod_{j=1}^{q} \frac{\Gamma_m(b_j)}{(b_j - m)!^m} \right] \left[\prod_{i=1}^{p} \frac{(a_i - m)!^m}{\Gamma_m(a_i)} \right] ,$$

$\tilde{a}_i = a_i - m + 1$, $i = 1, \ldots, p$, $\tilde{b}_j = b_j - m + 1$, $j = 1, \ldots, q$, and the eigenvalues of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are denoted by $\phi_1, \ldots, \phi_m$ and $\psi_1, \ldots, \psi_m$, respectively.

The ℓ-th derivative of the generalized hypergeometric function ${}_pF_q(\mathbf{a}; \mathbf{b}; sx)$ is given by [14]:

$$\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell} {}_pF_q(\mathbf{a}; \mathbf{b}; sx) = s^\ell \frac{(\mathbf{a})_\ell}{(\mathbf{b})_\ell} {}_pF_q(\tilde{\mathbf{a}}; \tilde{\mathbf{b}}; sx) \tag{7}$$

where $\tilde{a}_i = a_i + \ell$, $\tilde{b}_j = b_j + \ell$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, and s is a parameter.

5) Matrix spaces: We will denote by $\mathcal{U}(m)$ the unitary group of size m and by $\mathcal{S}(m, n)$ the Stiefel manifold of $m \times n$ matrices [6, Sec. II.C]. The region defined by the Stiefel manifold $\mathcal{S}(m, n)$, with $m \ge n$, is compact and has volume $|\mathcal{S}(m, n)| = 2^n \pi^{mn} / \Gamma_n(m)$. When $m = n$, the Stiefel manifold is a unitary group and its volume is given by $|\mathcal{U}(m)| = 2^m \pi^{m^2} / \Gamma_m(m)$.

B. Matrix-variate distributions

Definition 1: An $m \times n$ ($m \ge n$) random *Stiefel* matrix $\mathbf{S} \in \mathcal{S}(m, n)$ is such that $\mathbf{S}^{\mathsf{H}}\mathbf{S} = \mathbf{I}$ and is uniformly distributed on $\mathcal{S}(m, n)$. Then, it has pdf $p(\mathbf{S}) = |\mathcal{S}(m, n)|^{-1}$.

Definition 2: A square $m \times m$ random unitary matrix $\mathbf{U} \in \mathcal{U}(m)$ is such that $\mathbf{U}\mathbf{U}^{\mathsf{H}} = \mathbf{U}^{\mathsf{H}}\mathbf{U} = \mathbf{I}$. When it is uniformly distributed on $\mathcal{U}(m)$, it has pdf $p(\mathbf{U}) = |\mathcal{U}(m)|^{-1}$.

Definition 3: [15, Definition 2.6 and Lemma 2.6] An $m \times m$ Hermitian random matrix $\mathbf{A}$ is unitarily invariant if the joint distribution of its entries equals that of $\mathbf{V}\mathbf{A}\mathbf{V}^{\mathsf{H}}$ where $\mathbf{V}$ is any unitary matrix independent of $\mathbf{A}$. If $\mathbf{A}$ is unitarily invariant, then its eigenvalue decomposition can be written as $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ where $\mathbf{U}$ is a Haar matrix independent of the diagonal matrix $\mathbf{\Lambda}$. Since $\mathbf{U}$ is Haar (isotropic), it is uniformly distributed on $\mathcal{U}(m)$.

Definition 4: Let $\mathbf{H}$ be an $m \times n$ matrix whose columns are zero-mean independent complex Gaussian vectors with covariance matrix $\mathbf{\Theta}$.

- For $m \leq n$, the $m \times m$ random matrix $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is a central complex Wishart matrix, with n degrees of freedom and covariance matrix $\mathbf{\Theta}$ ($\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{\Theta})$). The joint distribution of the eigenvalues of $\mathbf{W}$ coincides with the law of the squared non-zero singular values of $\mathbf{H}$. Let $\mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ be the singular value decomposition (SVD) of $\mathbf{W}$. If $\mathbf{\Theta} = \mathbf{I}$, then $\mathbf{W}$ is unitarily invariant [15]. In such a case, the joint distribution of the ordered eigenvalues $\mathbf{\Lambda}$ can be written as [3], [11]
$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 |\mathbf{\Lambda}|^{n-m} \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\Gamma_m(n)\Gamma_m(m)} \mathcal{V}^2(\mathbf{\Lambda}) \,. \tag{8}$$
- For $m > n$, if the rows of $\mathbf{H}$ are independent and their covariance matrix is $\mathbf{I}$, the distribution of the ordered eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by [3]
$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 |\mathbf{\Lambda}|^{m-n} \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\Gamma_n(m)\Gamma_n(n)} \mathcal{V}^2(\mathbf{\Lambda}) \,. \tag{9}$$

Definition 5: Let $\mathbf{H}$ be an $m \times n$ random matrix whose entries are independent, complex, Gaussian random variables with unit variance and average $\mathbf{M} = \mathbb{E}[\mathbf{H}]$. Then, matrix $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is non-central Wishart [11].

- For $m \leq n$, the distribution of $\mathbf{W}$ is given by [11, eq. (99)]
$$p(\mathbf{W}) = \frac{|\mathbf{W}|^{n-m}}{\Gamma_m(n)} \frac{{}_0F_1(\,;n;\mathbf{M}\mathbf{M}^{\mathsf{H}}\mathbf{W})}{\mathrm{e}^{\mathsf{Tr}\{\mathbf{W}+\mathbf{M}\mathbf{M}^{\mathsf{H}}\}}} \,. \tag{10}$$
If $\mathbf{M}\mathbf{M}^{\mathsf{H}}$ has full rank and distinct eigenvalues, $\mu_1, \ldots, \mu_m$, then the joint pdf of the ordered, strictly positive eigenvalues $(\lambda_1, \ldots, \lambda_m) = \mathrm{diag}(\mathbf{\Lambda})$ of $\mathbf{W}$ is given by [11, eq. (102)]
$$p(\mathbf{\Lambda}) = \frac{|\mathbf{\Lambda}|^{n-m} \mathcal{V}(\mathbf{\Lambda}) |\{{}_0F_1(\,;n-m+1;\mu_i\lambda_j)\}|}{(n-m)!^m \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}+\mathbf{M}\mathbf{M}^{\mathsf{H}}\}} \mathcal{V}(\mathbf{M}\mathbf{M}^{\mathsf{H}})} \,. \tag{11}$$
Note that (11) has been obtained from [11, eq. (102)] by exploiting the result in (6).
As can be observed from (10), if $\mathbf{M}\mathbf{M}^{\mathsf{H}}$ is a scalar matrix (i.e., $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$), $p(\mathbf{W})$ only depends on the eigenvalues of $\mathbf{W}$. Thus $\mathbf{W}$ is unitarily invariant. In such a case, the distribution of $\mathbf{\Lambda}$ can be obtained from (11) by applying the limit in (3) and the property in (7), and it is given by
$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 |\mathbf{\Lambda}|^n |\mathbf{F}| \mathcal{V}(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n) \mathrm{e}^{\mu m + \mathsf{Tr}\{\mathbf{\Lambda}\}}} \tag{12}$$
where $(\mathbf{F})_{ij} = \lambda_j^{-i} {}_0F_1(\,;n-i+1;\mu\lambda_j)$.
Note that, since the Vandermonde determinant in (2) and pdf are positive by definition, here and in the following $|\mathbf{F}|$ represents the absolute value of the determinant of matrix $\mathbf{F}$. This avoids us to include in the provided results coefficients that account for the sign of determinants.
- For $m > n$, the same expressions as in (10), (11), and (12) hold but replacing, $\mathbf{M}$, m and n with, respectively, $\mathbf{M}^{\mathsf{H}}$, n and m.

Lemma 1: Let $\mathbf{H}_1$ and $\mathbf{H}_2$ be, respectively, an $m \times n$ and an $m \times p$ ($m \leq p$) Gaussian complex random matrix whose columns are independent, have zero mean, and covariance $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, respectively.

- For $m \leq n$, the $m \times m$ random matrix $\mathbf{W} = (\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ is a central $\mathcal{F}$-matrix [11]. When $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are both scalar matrices, $\mathbf{W}$ is unitarily invariant and has a Beta type II distribution [16]. Specifically, when $\mathbf{\Theta}_1\mathbf{\Theta}_2^{-1} = \omega\mathbf{I}$, the distribution of its ordered eigenvalues is given by
$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 \Gamma_m(p+n)}{\omega^{mn}\Gamma_m(m)\Gamma_m(p)\Gamma_m(n)} \frac{\mathcal{V}^2(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{|\mathbf{I}+\mathbf{\Lambda}/\omega|^{p+n}} \,. \tag{13}$$
- For $m > n$, the matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is unitarily invariant and the distribution of its ordered eigenvalues can be expressed as
$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 \Gamma_m(p+n)|\mathbf{F}||\mathbf{\Omega}|^{-n}\mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}|^{m-p-n-1}}{(p+n-m)!^{-n}\Gamma_n(p+n)\Gamma_m(p)\Gamma_n(n)\mathcal{V}(\mathbf{I}-\mathbf{\Omega}^{-1})} \tag{14}$$
where $\mathbf{\Omega} = \mathbf{\Theta}_1^{1/2}\mathbf{\Theta}_2^{-1}\mathbf{\Theta}_1^{1/2}$, $(\mathbf{F})_{ij} = {}_1F_0(p+n-m+1;\,;(1-\omega_i^{-1})\lambda_j/(1+\lambda_j))$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$, and $(\mathbf{F})_{ij} = (1-\omega_i^{-1})^{m-j}$ for $i = 1, \ldots, m$, $j = n+1, \ldots, m$.

Proof: The proof is given in Appendix A. ∎

Lemma 2: Let $\mathbf{H}_1$ and $\mathbf{H}_2$ be, respectively, an $m \times n$ and an $m \times p$ Gaussian complex random matrix whose columns are independent and have covariance $\mathbf{\Theta}$. Let also $\mathbb{E}[\mathbf{H}_1] = \mathbf{M}$ and $\mathbb{E}[\mathbf{H}_2] = \mathbf{0}$.

- For $m \leq n$, $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$ and $\mathbf{\Theta} = \theta\mathbf{I}$, the non-central $\mathcal{F}$-matrix $\mathbf{W} = (\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ is unitarily invariant and the distribution of its eigenvalues is given by
$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 \mathcal{V}(\mathbf{\Lambda})^2 \Gamma_m(p+n) {}_1F_1(p+n;n;\frac{\mu}{\theta}\mathbf{\Lambda}(\mathbf{I}+\mathbf{\Lambda})^{-1})}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(p)\mathrm{e}^{\mu m/\theta}|\mathbf{\Lambda}|^{m-n}|\mathbf{I}+\mathbf{\Lambda}|^{p+n}} \tag{15}$$
- For $m > n$, and $\mathbf{M}^{\mathsf{H}}\mathbf{\Theta}^{-1}\mathbf{M} = \omega\mathbf{I}$, the matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is unitarily invariant and the distribution of its eigenvalues is given by
$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 \Gamma_n(p+n)\mathrm{e}^{-\omega n}}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(p+n-m)} \frac{|\mathbf{F}||\mathbf{\Lambda}|^{m-n}\mathcal{V}(\mathbf{\Lambda})}{|\mathbf{I}+\mathbf{\Lambda}|^{p+1}} \tag{16}$$
where $(\mathbf{F})_{ij} = (\lambda_j/(1+\lambda_j))^{n-i} {}_1F_1(p+n-i+1;m-i+1;\omega\lambda_j/(1+\lambda_j))$.

The proof is provided in Appendix B.

Definition 6: The $n \times n$ random matrix $\mathbf{B}$ is Beta-distributed with positive integer parameters p and q ($\mathbf{B} \sim \mathcal{B}_n(p,q)$) if

- given $\mathbf{T}$ an upper triangular matrix with positive diagonal elements, we can write $\mathbf{B} = (\mathbf{T}^{\mathsf{H}})^{-1}\mathbf{C}\mathbf{T}$ where $\mathbf{C} \sim \mathcal{W}_n(p, \mathbf{\Theta})$, and
- given $\mathbf{A} \sim \mathcal{W}_n(m, \mathbf{\Theta})$, we can write $\mathbf{A} + \mathbf{C} = \mathbf{T}^{\mathsf{H}}\mathbf{T}$. Notice that, if either $p < n$ or $q < n$, or both $p < n$ and $q < n$, the distribution is referred to as *pseudo*-Beta since it involves *pseudo*-Wishart matrices [5, and references therein].

When $n \leq p$, $\mathbf{B}$ admits an eigendecomposition where the matrix of the eigenvectors is independent of the matrix of the eigenvalues [5, Lemma 8].

- For $q \leq n$, the distribution of the q ordered non-zero eigenvalues of $\mathbf{B}$ is given by [5, eq. (13)]:

$$p(\mathbf{\Lambda}) = \frac{\pi_q^2 \Gamma_q(p+q)|\mathbf{I} - \mathbf{\Lambda}|^{n-q}|\mathbf{\Lambda}|^{p-n}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_q(n)\Gamma_q(p+q-n)\Gamma_q(q)}. \quad (17)$$

- For $q > n$, $\mathbf{B}$ has n nonzero eigenvalues, whose ordered joint distribution is given by [5, eq. (12)]:

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 \Gamma_n(p+q)|\mathbf{I} - \mathbf{\Lambda}|^{q-n}|\mathbf{\Lambda}|^{p-n}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_n(n)\Gamma_n(p)\Gamma_n(q)}. \quad (18)$$

Due to the lack of the corresponding expression in the literature, herein we derive the expression of the marginal distribution of a single unordered eigenvalue of a $\mathcal{B}_n(p,q)$-distributed matrix, which will be needed in our subsequent derivations.

Proposition 1: Given an $n \times n$ matrix $\mathbf{B} \sim \mathcal{B}_n(p,q)$,

- For $q \leq n$, the pdf of a single unordered eigenvalue of $\mathbf{B}$ is given by

$$p(\lambda) = \frac{\pi_q^2}{q\Gamma_q(q)} \frac{\Gamma_q(p+q)\Gamma(n-q+1)}{\Gamma_q(n)\Gamma_q(p+q-n)} \cdot \sum_{i,j=1}^{n} \lambda^{p-n+i+j-2}(1-\lambda)^{n-q}\mathcal{D}_{ij} \quad (19)$$

  with $\mathcal{D}_{ij}$ being the (i,j)-cofactor of the $(n \times n)$ matrix $\mathbf{A}$ such that

$$(\mathbf{A})_{\ell k} = \frac{\Gamma(p-n+\ell+k-1)}{\Gamma(p+k-q+\ell)}. \quad (20)$$

- For $q > n$, the pdf of a single unordered eigenvalue of $\mathbf{B}$ is given by

$$p(\lambda) = \frac{\pi_n^2}{n\Gamma_n(n)} \frac{\Gamma_n(p+q)\Gamma(q-n+1)}{\Gamma_n(p)\Gamma_n(q)} \cdot \sum_{i,j=1}^{n} \lambda^{(p-n+i+j-2)}(1-\lambda)^{q-n}\mathcal{D}_{ij} \quad (21)$$

  with $\mathcal{D}_{ij}$ being the (i,j)-cofactor of the $(n \times n)$ matrix $\mathbf{A}$ such that

$$(\mathbf{A})_{\ell k} = \frac{\Gamma(p-n+\ell+k-1)}{\Gamma(p+k+q-2n+\ell)}. \quad (22)$$

Proof: The proof is given in Appendix C. ■

C. Communication model

We consider a single-user multiple-antenna communication system, with m and n denoting the number of receive and transmit antennas, respectively. Assuming block-memoryless fading with coherence length equal to b, the output can be described by the following linear relationship:

$$\mathbf{Y} = \sqrt{\gamma}\mathbf{H}\mathbf{X} + \mathbf{N} \quad (23)$$

where $\mathbf{Y}$ is the $m \times b$ output matrix, and $\mathbf{H}$ is the $m \times n$ complex random channel matrix whose entries represent the fading coefficients between each transmit and receive antenna. $\mathbf{N}$ is the $m \times b$ matrix of white Gaussian noise which is assumed to have i.i.d. complex Gaussian entries with zero mean and unitary variance. The normalized per-transmit antenna SNR is denoted by $\gamma = \mathrm{SNR}/n$, and $\mathbf{X}$ is the random complex $n \times b$ input matrix whose structure will be specified in the following sections. Moreover, for any positive integer n, we define

$$\gamma_n = \gamma^{n(n-1)/2}.$$

Note that the above communication model is adopted in all the following sections, except for Section III-B where we resort to a slightly different model explicitly accounting for interference.

III. Output Statistics with IID Gaussian Input

In this section, we analyse the case where the distribution of $\mathbf{X}$ is Gaussian i.i.d. and consider both the noise-limited and interference-limited scenarios. Note that, in the case under study, the average energy of the input signal is given by $\mathbb{E}[\mathsf{Tr}\{\mathbf{X}\mathbf{X}^{\mathsf{H}}\}] = nb$.

As for the communication channel, we focus our analysis on some classes of channel matrices whose Gramian $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant. As shown in the following, this allows us to write the expression of the output pdf in terms of the distribution of the eigenvalues of the channel matrix. In particular, in both the noise-limited and the interference-limited case, we draw on the following results:

- for $m \leq n$, and for unitarily invariant $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, the distribution of $\mathbf{Y}$ is given by [3, eq. (40) and (41)]

$$p(\mathbf{Y}) = \frac{\Gamma_m(m)K(\mathbf{Y})}{\pi_m\gamma_m} \int \frac{|\mathbf{E}||\mathbf{I} + \gamma\mathbf{\Lambda}|^{m-b-1}}{\mathcal{V}(\mathbf{\Lambda})} p(\mathbf{\Lambda})\,\mathrm{d}\mathbf{\Lambda}, \quad (24)$$

  where $\mathbf{\Lambda}$ is an $m \times m$ diagonal matrix containing the eigenvalues of channel matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, $(\mathbf{E})_{ij} = \mathrm{e}^{y_i c_j}$, and $c_j = \gamma\lambda_j/(1+\gamma\lambda_j)$, $j = 1, \ldots, m$. Moreover, $y_1, \ldots, y_m$ are the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ and

$$K(\mathbf{Y}) = \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2}}{\pi^{mb}\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})}. \quad (25)$$

- for $m > n$, and for unitarily invariant $\mathbf{H}^{\mathsf{H}}\mathbf{H}$, the pdf of $\mathbf{Y}$ can be obtained by following the steps described in [3] and is given by

$$p(\mathbf{Y}) = \frac{\Gamma_n(m)K(\mathbf{Y})}{\pi_n\gamma^{n(m-n)}} \int \frac{|\widetilde{\mathbf{E}}||\mathbf{I} + \gamma\mathbf{\Lambda}|^{m-b-1}}{\mathcal{V}(\gamma\mathbf{\Lambda})|\mathbf{\Lambda}|^{m-n}} p(\mathbf{\Lambda})\,\mathrm{d}\mathbf{\Lambda} \quad (26)$$

where $\widetilde{\mathbf{E}}$ is an $m \times m$ matrix whose elements are given by $(\widetilde{\mathbf{E}})_{ij} = \mathrm{e}^{y_i c_j}$ for $1 \le j \le n$, and $(\widetilde{\mathbf{E}})_{ij} = y_i^{j-n-1}$ for $n+1 \le j \le m$. Note that in this case the matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is of reduced rank since it has $m-n$ zero eigenvalues. Thus, here $p(\mathbf{\Lambda})$ indicates the distribution of the n non-zero eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ and $\mathbf{\Lambda}$ is an $n \times n$ diagonal matrix.

Proof: The proof is given in Appendix D. ∎

A. Noise-limited

The output pdf of the uncorrelated Rayleigh-faded channel has been evaluated in [3]. For sake of completeness, we recall this result and present the corrected expression of the output pdf when $m > n$. Then, we extend the analysis to two other practically relevant fading models, namely, the Rician block-fading channel [17], [18] and Land Mobile Satellite (LMS) channel [9], [19].

1) Rayleigh fading channel: In the case of uncorrelated Rayleigh channel, the entries of $\mathbf{H}$ follow an i.i.d. zero-mean, unit-variance, complex Gaussian distribution.

- For $m \le n$, the distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is given by (8). It follows that, by using (24) and the result in Appendix K, the distribution of $\mathbf{Y}$ can be written as [3, Proposition 2].

$$p(\mathbf{Y}) = \frac{\pi_m}{\gamma_m \Gamma_m(n)} K(\mathbf{Y}) |\mathbf{Z}| \tag{27}$$

where the i,j-th entry of the $m \times m$ matrix $\mathbf{Z}$ is given by

$$(\mathbf{Z})_{ij} = \int_0^\infty \exp\left(\frac{y_i \gamma x}{1+\gamma x} - x\right) \frac{x^{n-m+j-1}}{(1+\gamma x)^{b+1-m}} \,\mathrm{d}x\,.$$

- For $m > n$, the distribution of the eigenvalues of channel matrix $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by (9). By applying (26) and the result in Appendix K, the output pdf is given by

$$p(\mathbf{Y}) = \frac{\pi_n}{\Gamma_n(n)\gamma^{n(m-n)}} K(\mathbf{Y}) |\mathbf{Z}|\,. \tag{28}$$

Note that the expression above differs from the one presented in [3, Proposition 2] in the term $\gamma^{n(m-n)}$, which appears at the denominator. The i,j-th entry of the $m \times m$ matrix $\mathbf{Z}$ can be written as

$$(\mathbf{Z})_{ij} = \int_0^\infty \exp\left(\frac{y_i \gamma x}{1+\gamma x} - x\right) \frac{(x/\gamma)^{j-1}}{(1+\gamma x)^{b+1-m}} \,\mathrm{d}x\,,$$

for $1 \le i \le m, 1 \le j \le n$ and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \le i \le m, n+1 \le j \le m$.

2) Rician channel: The Rician channel is traditionally modeled as a superposition of a scattered plus a LOS component, i.e.,

$$\mathbf{H} = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}} + \sqrt{\frac{1}{\kappa+1}}\widetilde{\mathbf{H}}\,. \tag{29}$$

In (29), κ is the Rician factor representing the ratio of the average power of the unfaded channel component to the faded channel component, the entries of $\widetilde{\mathbf{H}}$ are independent, zero-mean unit-variance complex Gaussian, and $\bar{\mathbf{H}}$ is a deterministic matrix representing the LOS component.

Specifically, for $m \le n$, we consider the special case $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$ (for $m > n$ we assume $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$), where h is a positive parameter. This assumption reflects two main settings: the scalar LOS channel, introduced in [7] and therein already analysed in the high-SNR regime, and the LOS MIMO with residual scattering [8]. Both models assume the LOS matrix to have high (full) rank. The one in [8] is suitable for MIMO backhaul links where antenna spacing is carefully designed and transmit-receive distance is fixed. Our model can be thought of as a Gaussian perturbation, with small variance, of the one in [8]. The model in [7], although being a sub-case of the one in [8] from the pure mathematical viewpoint, has played a major role in the early characterization of MIMO Rician channels, due to the amenability of diagonal [20] (and, in particular, scalar) non-centrality matrices for the derivation of the capacity-achieving input law.

Under the aforementioned assumption, the Gramian of the matrix $\mathbf{H}$ is unitarily invariant (see Definition 5), thus the pdf of the output can be expressed as in the following proposition.

Proposition 2: Given a channel as in (23) and (29), with i.i.d. Gaussian input and Rician block-fading,

- for $m \le n$, and $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$, the pdf of its output can be written as

$$p(\mathbf{Y}) = \frac{\pi_m (1+\kappa)^{mn}}{\gamma_m \Gamma_m(n) \mathrm{e}^{\kappa h m}} K(\mathbf{Y}) |\mathbf{Z}|\,, \tag{30}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \gamma x/(1+\gamma x)}\, {}_0F_1(\,;n-j+1;\hat{x})\,\mathrm{d}x}{\mathrm{e}^{(1+\kappa)x} x^{j-n} (1+\gamma x)^{b-m+1}}$$

- for $m > n$, and $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$, the following result holds

$$p(\mathbf{Y}) = \frac{\pi_n (1+\kappa)^{nm}}{\gamma_n \gamma^{n(m-n)} \Gamma_n(n) \mathrm{e}^{\kappa h n}} K(\mathbf{Y}) |\mathbf{Z}| \tag{31}$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \gamma x/(1+\gamma x)}\, {}_0F_1(\,;m-j+1;\hat{x})\,\mathrm{d}x}{\mathrm{e}^{(1+\kappa)x} x^{j-n} (1+\gamma x)^{b-m+1}}\,,$$

for $1 \le i \le m, 1 \le j \le n$ and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \le i \le m, n+1 \le j \le m$,

with $\hat{x} = \kappa(1+\kappa)hx$.

Proof: The proof is given in Appendix E. ∎

3) Land mobile satellite communication: The Land Mobile Satellite (LMS) MIMO channel can be viewed as a non-central channel with random mean. Thus, the channel matrix model can be described as

$$\mathbf{H} = \bar{\mathbf{H}} + \widetilde{\mathbf{H}} \tag{32}$$

where the entries of $\widetilde{\mathbf{H}}$ are independent, zero-mean unit-variance complex Gaussian and $\bar{\mathbf{H}}$ is a random matrix. As shown in [9], in a LMS channel the matrix $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}}$ follows a matrix-variate $\Gamma(\alpha, \mathbf{\Omega})$ distribution [21] where α plays the role of a shape parameter, while $\mathbf{\Omega}$ is a scale parameter. Indeed, α can be viewed as a generalized number of degrees of freedom of the non-centrality parameter, while $\mathbf{\Omega}$ is related to the average power of the random LOS component, as discussed in detail in [9]. Assuming $\mathbf{\Omega} = \omega\mathbf{I}$, $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant, as shown in [9, Property 1].

Under this assumption, the expression of the output pdf can be expressed as in the following proposition.

Proposition 3: Given an LMS MIMO channel as in (32) with $\mathbf{\Omega} = \omega \mathbf{I}$,

- for $m \leq n$, the pdf of its output can be written as

$$p(\mathbf{Y}) = \frac{\pi_m}{\gamma_m \Gamma_m(n)(1+1/\omega)^{m\alpha}} K(\mathbf{Y}) |\mathbf{Z}|\,, \qquad (33)$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}} {}_1F_1\left(\alpha - j + 1; n - j + 1; \frac{x}{1+\omega}\right)}{\mathrm{e}^x x^{j-n}(1+\gamma x)^{b-m+1}} \,\mathrm{d}x\,;$$

- for $m > n$, the output pdf is given by:

$$p(\mathbf{Y}) = \frac{\pi_n}{\gamma_n \gamma^{n(m-n)} \Gamma_n(m)(1+1/\omega)^{n\alpha}} K(\mathbf{Y}) |\mathbf{Z}| \quad (34)$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}} {}_1F_1\left(\alpha - j + 1; m - j + 1; \frac{x}{1+\omega}\right)}{\mathrm{e}^x x^{j-n}(1+\gamma x)^{b-m+1}} \,\mathrm{d}x\,,$$

for $1 \leq i \leq m, 1 \leq j \leq n$ and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \leq i \leq m, n+1 \leq j \leq m$.

Proof: The proof is given in Appendix F. ∎

B. Interference limited

We now consider the case where the main impairment to communication is represented by the co-channel interference. In particular, each interferer is seen from the direct link receiver under its own random channel, which we assume to be affected by Rayleigh fading, again with block-length b. We assume that there are L active interferers in the network, each equipped, for homogeneity, with the same number of antennas, n, as the transmitter of the useful signal. We evaluate the output pdf when a whitening filter is applied to the received signal and we consider two channel models. In the former, the desired signal undergoes Rayleigh fading; in the latter, the direct link is affected by Rician fading, i.e., we assume the existence of an LOS path between the useful transmitter and its intended receiver.

The received signal can be modeled as

$$\widetilde{\mathbf{Y}} = \sqrt{\gamma} \mathbf{H}_s \mathbf{X} + \mathbf{W} \qquad (35)$$

where

$$\mathbf{W} = \sum_{\ell=1}^{L} \widehat{\mathbf{H}}_\ell \widehat{\mathbf{X}}_\ell$$

represents the interference. Specifically, the $m \times n$ matrix $\widehat{\mathbf{H}}_\ell$ models the channel connecting the ℓ-th interferer with the receiver, while the $n \times b$ matrix $\widehat{\mathbf{X}}_\ell$ represents the signal transmitted by the ℓ-th interferer, $\ell = 1, \ldots, L$. The interference can be rewritten as $\mathbf{W} = \widehat{\mathbf{H}}\widehat{\mathbf{X}}$ where $\widehat{\mathbf{H}} = [\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_L]$ is an $m \times Ln$ matrix and $\widehat{\mathbf{X}} = [\widehat{\mathbf{X}}_1{}^{\mathsf{H}}, \ldots, \widehat{\mathbf{X}}_L{}^{\mathsf{H}}]^{\mathsf{H}}$ is of size $Ln \times b$. By assuming that the entries of $\widehat{\mathbf{X}}$ are i.i.d. complex Gaussian with zero mean and unit variance, the covariance of the interference, conditioned on the knowledge of the composite channel matrix $\widehat{\mathbf{H}}$, is given by

$$\mathbf{R} = \mathbb{E}[\mathbf{W}\mathbf{W}^{\mathsf{H}} | \widehat{\mathbf{H}}] = \widehat{\mathbf{H}} \mathbb{E}[\widehat{\mathbf{X}}\widehat{\mathbf{X}}^{\mathsf{H}}] \widehat{\mathbf{H}}^{\mathsf{H}} = b \widehat{\mathbf{H}} \widehat{\mathbf{H}}^{\mathsf{H}}\,.$$

We apply to the received signal $\widetilde{\mathbf{Y}}$ the whitening filter $\mathbf{B} = \sqrt{b}\mathbf{R}^{-1/2}$ and obtain

$$\begin{aligned} \mathbf{Y} &= \mathbf{B}\widetilde{\mathbf{Y}} \\ &= \sqrt{b}\mathbf{R}^{-1/2}\widetilde{\mathbf{Y}} \\ &= \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2} (\sqrt{\gamma}\mathbf{H}_s\mathbf{X} + \mathbf{W}) \\ &= \sqrt{\gamma}\mathbf{H}\mathbf{X} + \mathbf{N} \end{aligned} \qquad (36)$$

where $\mathbf{H} = \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2} \mathbf{H}_s$ and $\mathbf{N} = \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2} \mathbf{W}$. Clearly, $\mathbb{E}[\mathbf{N}\mathbf{N}^{\mathsf{H}} | \widehat{\mathbf{H}}] = b\mathbf{I}$. In the following, we provide the pdf of $\mathbf{Y}$.

1) Rayleigh fading channel:

Proposition 4: We consider the interference-limited channel described by (35), with L active interferers, i.i.d. Gaussian input and Rayleigh fading. If $\mathbf{H}_s\mathbf{H}_s{}^{\mathsf{H}} \sim \mathcal{W}_m(n, \mathbf{\Theta}_s)$ and $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}} \sim \mathcal{W}_m(Ln, \widehat{\mathbf{\Theta}})$, then we have the following results.

- For $m \leq n$, due to mathematical constraints, we only analyse the case of spatially uncorrelated receiving antennas, i.e., $\mathbf{\Theta}_s = \theta_s \mathbf{I}$ and $\widehat{\mathbf{\Theta}} = \hat{\theta}\mathbf{I}$. Then, the pdf of $\mathbf{Y}$ can be written as

$$p(\mathbf{Y}) = \frac{\pi_m \Gamma_m(Ln+n)}{\gamma_m \omega^{mn} \Gamma_m(Ln)\Gamma_m(n)} K(\mathbf{Y}) |\mathbf{Z}| \qquad (37)$$

where $\omega = \theta_s/\hat{\theta}$ and

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}} x^{n-j}}{(1+\gamma x)^{b-m+1}(1+x/\omega)^{Ln+n}} \,\mathrm{d}x\,.$$

This result is obtained by substituting (13) in (24) and by exploiting the result in Appendix K.

- For $m > n$, the pdf of $\mathbf{Y}$ is given by:

$$\begin{aligned} p(\mathbf{Y}) \;=\; & \frac{\pi_n (Ln+n-m)!^m \Gamma_n(m)\Gamma_m(Ln+n) K(\mathbf{Y})}{\Gamma_n(Ln+n)\Gamma_n(n)\Gamma_m(Ln)\gamma_n\gamma^{n(m-n)}} \\ & \cdot \frac{|\mathbf{\Omega}|^{m-n-1}}{\mathcal{V}(\mathbf{\Omega})} \int \frac{|\widetilde{\mathbf{E}}||\mathbf{I}+\gamma\mathbf{\Lambda}|^{m-b-1}|\mathbf{F}|\,\mathrm{d}\mathbf{\Lambda}}{|\mathbf{\Lambda}|^{m-n}|\mathbf{I}+\mathbf{\Lambda}|^{Ln+n-m+1}} \end{aligned} \qquad (38)$$

where $\mathbf{\Omega} = \mathbf{\Theta}_1^{1/2}\mathbf{\Theta}_2^{-1}\mathbf{\Theta}_1^{1/2}$ and the matrices $\widetilde{\mathbf{E}}$ and $\mathbf{F}$ have been defined below (26) and (14), respectively. This result is obtained by substituting (14) in (26). However, we cannot solve the integral by applying the result in Appendix K directly. Indeed, although matrices $\widetilde{\mathbf{E}}$ and $\mathbf{F}$ are both of size $m \times m$, a portion of their columns and rows is composed of constant terms. Thus, we need to resort to the property of the determinant of block matrices, in order to obtain $n \times n$ blocks to which the result in Appendix K can be applied. We skip the details of this procedure due to the cumbersome expressions that are involved.

2) Rician fading channel:

Proposition 5: We consider the interference-limited channel described by (35), with L active interferers, i.i.d. Gaussian input, Rician faded useful signal and Rayleigh fading affecting the interfering links. For a Rician channel, matrix $\mathbf{H}_s$ can be written as in (29)

$$\mathbf{H}_s = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}}_s + \sqrt{\frac{1}{\kappa+1}}\widetilde{\mathbf{H}}_s$$

where κ is the Rician factor, $\bar{\mathbf{H}}_s$ is deterministic, and $\widetilde{\mathbf{H}}_s$ is complex Gaussian with independent colums whose covariance is $\mathbf{\Theta}$. According to our assumptions on LOS links made in Section III-A, we have:

- for $m \leq n$, setting $\mathbf{\Theta}_s = \widehat{\mathbf{\Theta}} = \theta\mathbf{I}$ and $\bar{\mathbf{H}}_s\bar{\mathbf{H}}_s{}^{\mathsf{H}} = h\mathbf{I}$,

$$p(\mathbf{Y}) = \frac{\pi_m \Gamma_m(Ln+n)\mathrm{e}^{-h\kappa m/\theta}}{\gamma_m \Gamma_m(n)\Gamma_m(Ln)\tilde{\kappa}^{-mn}} K(\mathbf{Y})|\mathbf{Z}| \qquad (39)$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{\mathrm{e}^{y_i \frac{\gamma x}{1+\gamma x}}{}_1F_1(\tilde{L}+j; n-m+j; h\kappa\tilde{\kappa}\tilde{x}/\theta)\,\mathrm{d}x}{(1+\gamma x)^{b-m+1}(1+\tilde{\kappa}x)^{\tilde{L}+1}x^{m-n}\tilde{x}^{1-j}} \qquad (40)$$

with $\tilde{\kappa} = 1+\kappa$, $\tilde{L} = Ln+n-m$, and $\tilde{x} = x/(1+\tilde{\kappa}x)$

- for $m > n$, and $\bar{\mathbf{H}}_s{}^{\mathsf{H}}\mathbf{\Theta}^{-1}\bar{\mathbf{H}}_s = h\mathbf{I}$,

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_n(Ln+n)\tilde{\kappa}^{nm}\mathrm{e}^{-h\kappa n}}{\gamma_n \gamma^{n(m-n)}\Gamma_n(n)\Gamma_n(Ln+n-m)} K(\mathbf{Y})|\mathbf{Z}| \qquad (41)$$

where

$$(\mathbf{Z})_{ij} = \int_0^\infty \frac{{}_1F_1(Ln+n-j+1; m-j+1; h\kappa\tilde{\kappa}\tilde{x})\,\mathrm{d}x}{\mathrm{e}^{\frac{-y_i\gamma x}{1+\gamma x}}(1+\tilde{\kappa}x)^{Ln+1}\tilde{x}^{j-n}(1+\gamma x)^{b-m+1}},$$

for $1 \leq i \leq m, 1 \leq j \leq n$, and $(\mathbf{Z})_{ij} = y_i^{j-n-1}$, for $1 \leq i \leq m, n+1 \leq j \leq m$; with $\tilde{\kappa} = 1+\kappa$ and $\tilde{x} = x/(1+\tilde{\kappa}x)$.

Proof: The proof is given in Appendix G. ∎

Note that also in this case mathematical issues made the analysis only possible for uncorrelated receivers.

C. Exploitation of the analytical results

The mutual information between the channel input, $\mathbf{X}$, and the channel output, $\mathbf{Y}$, normalized to the fading coherence length, can be expressed as:

$$\mathcal{I} = \frac{1}{b}\left[h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X})\right] \qquad (42)$$

where $h(\mathbf{Y}) = \mathbb{E}[-\log p(\mathbf{Y})]$ and $h(\mathbf{Y}|\mathbf{X}) = \mathbb{E}[-\log p(\mathbf{Y}|\mathbf{X})]$. Once the pdf of the channel output, $p(\mathbf{Y})$, is obtained, it can be used to evaluate its differential entropy, $h(\mathbf{Y})$. For Rayleigh and Gaussian channels with identity covariance matrix, considering that $\mathbf{X}$ is given, the output $\mathbf{Y}$ is complex Gaussian and its rows are i.i.d. Hence, in order to derive the conditional differential entropy $h(\mathbf{Y}|\mathbf{X})$, we can compute its value for an arbitrary row of $\mathbf{Y}$ and then scale it by the number of rows of $\mathbf{Y}$ [3].

In [3], the mutual information has been computed in presence of Rayleigh channel and i.i.d. Gaussian input, for $m \leq n$. In the following, we provide three examples of mutual information computation. First, we address the case of noise-limited Rayleigh channel with $m > n$ and, then, the noise-limited Rician channel, both with $m \leq n$ and $m > n$.

In the case of Rayleigh channel, the conditional differential entropy is obtained using [3, eq. (4)], while the unconditional differential entropy is evaluated using (27) or (28) depending on the relationship between m and n. Fig. 1 shows the mutual information as a function of the SNR, with $b = 6, 10$, $m = 2$ and $n = 1$, when no channel state information (CSI) is available and in the case of perfect CSI at the receiver. The latter is obtained by computing [3, eq. (10)]. The results confirm the intuition, as well as previous analysis [4], [6]: the higher the SNR and the value of b, the better the performance, while the lack of CSI causes a noticeable degradation.

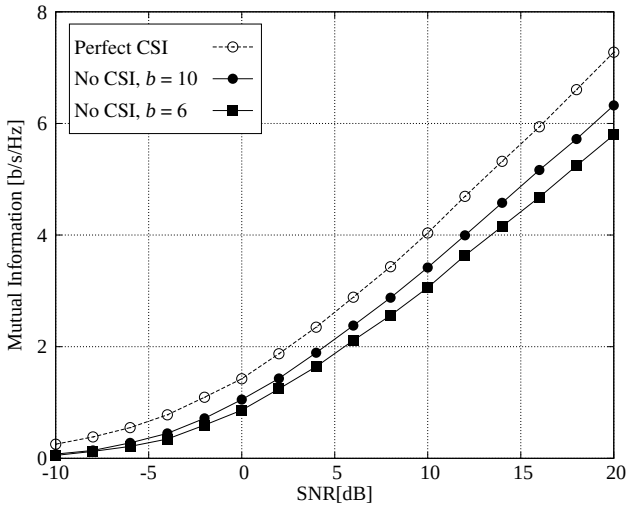

Fig. 1. Mutual information vs. SNR in Rayleigh channel: comparison between the case where no CSI is available (solid line) and the case of perfect CSI at the receiver (dashed line), with $b = 6, 10$, $m = 2$ and $n = 1$.

For the Rician channel, the expression of the channel matrix is given by (29). By adopting again the method in [3], the differential entropy of the output conditioned on the input signal can be computed. Let us denote by $\mathbf{y}$ an arbitrary row of $\mathbf{Y}$; then, using [3, eq. (31)] and considering the translation-invariant property of differential entropy, we can write the mutual information when the receiver does not have any knowledge of the non-LOS component:

$$h(\mathbf{y}|\mathbf{X}) = h(\mathbf{y}^{\mathsf{H}}|\mathbf{X}) = \mathbb{E}\left[\log_2\left((\pi\mathrm{e})^b\left|\mathbf{I}+\frac{\gamma\mathbf{X}^{\mathsf{H}}\mathbf{X}}{1+\kappa}\right|\right)\right] \qquad (43)$$

with the expectation being over the distribution of $\mathbf{X}$. The above expression can be conveniently computed resorting to [3, eq. (4)]. The unconditional differential entropy of the output is derived through (30) and (31).

Fig. 2 shows the mutual information as a function of the SNR, with $b = 6$, $m = 2$ and $n = 2$. Rician factors are set to $\kappa = 1$ and $\kappa = 10$. The plot depicts the mutual information in the two cases where the receiver has knowledge of the non-LOS component [3, eq. (10)] and where it does not (43). The deterministic channel matrix in (29) is set as follows:

$$\bar{\mathbf{H}} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}.$$

In Fig. 2, the relative gap between the achievable mutual information in the two scenarios with $\kappa = 1$ is more evident than for $\kappa = 10$, since the higher the Rician factor, the higher the amount of information on the LOS component, which is known at the receiver. This is also compliant with the monotonicity results in [20].

Finally, Fig. 3 shows the mutual information for the two scenarios above, in the case of $m > n$, namely, $m = 2$, $n =$

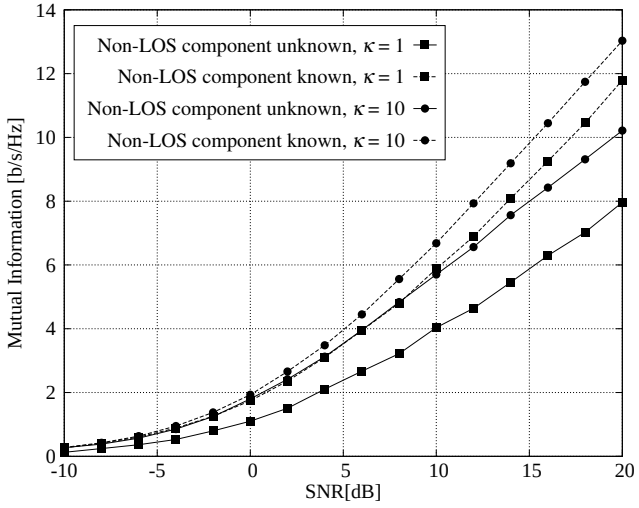

Fig. 2. Mutual information vs. SNR in Rician channel: comparison between the case where the receiver does not have any knowledge on the non-LOS component (solid line) and when such knowledge is available (dashed line), for $b = 6$, $n = 2$, $m = 2$ and $\kappa = 1, 10$.

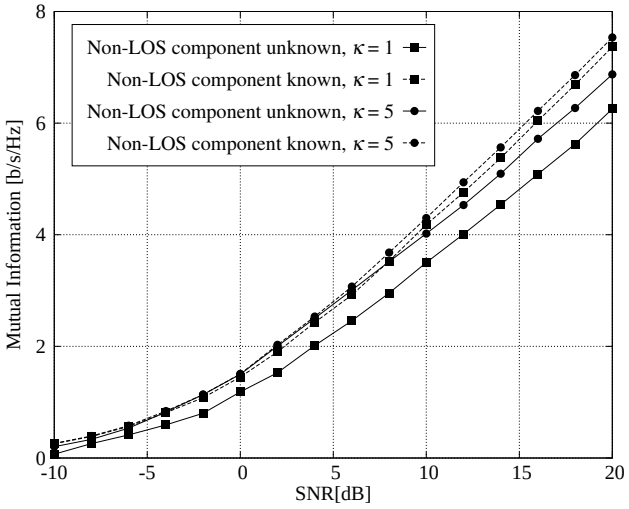

Fig. 3. Mutual information vs. SNR in Rician channel: comparison between the cases where knowledge of the non-LOS component is not available at the receiver (solid line) and when it is (dashed line). $b = 6$, $m = 2$, $n = 1$ and $\kappa = 1, 5$.

1, and $b = 6$. The Rician factor is set to $\kappa = 1$ and $\kappa = 5$. In this scenario, the deterministic channel matrix is set to $\bar{\mathbf{H}} = \left[\sqrt{3/2}, 1/\sqrt{2}\right]^H$. Similar observations to those above hold. However, comparing Fig. 2 to Fig. 3, we notice that, as expected, the reduction in the number of antennas at the transmitter leads to severe performance degradation.

IV. Output Statistical Characterization with Product Input Form

As in [3], [6], [22], [23], we assume total lack of CSI at both the ends of the wireless link. This case is of particular interest for the energy efficiency of the communication, as the availability of CSI would imply a high energy and time consumption at both the transmitter and the receiver. Under this assumption, in the high-SNR regime, the capacity-achieving input matrix $\mathbf{X}$ is proven to have a product structure [22, theorem 2] and can be written as

$$\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi} \tag{44}$$

where c is a normalizing constant and $\mathbf{D}$ is a real random $n \times n$ diagonal matrix, which is positive definite with probability 1. The entries of $\mathbf{D}$ represent the amount of transmit power allocated to each of the n transmit antennas, while $\mathbf{\Phi} \in \mathcal{S}(n, b)$ represents the beamforming $n \times b$ matrix. In order to be consistent with the definition of SNR, we impose the constraint on the average input energy $\mathbb{E}[\mathsf{Tr}\{\mathbf{X}\mathbf{X}^{\mathsf{H}}\}] = nb$. It follows that, for our specific input structure, the normalizing constant is given by:

$$c = \frac{nb}{\mathbb{E}[\mathsf{Tr}\{\mathbf{D}\}]}\,. \tag{45}$$

In [5, Lemma 10], it is proven that without CSI, in Rayleigh block-fading channels, the optimal power allocation at the transmitter depends on the relationship between the coherence length, b, and the total number of antennas at both the transmitter and receiver. Specifically,

- If $b \geq m+n$, all diagonal entries of $\mathbf{D}$ are almost surely equal to 1. This case corresponds to the conventional unitary space-time modulation (USTM) [22], where $\mathbf{D} = \mathbf{I}$ and $c = b$;
- If $b < m + n$, the optimal input is $\mathbf{D} \sim \mathcal{B}_n(b - n, m + n - b)$, which is referred to as Beta-variate space-time modulation (BSTM) [5]. This scenario allows the analysis of an uplink massive-MIMO system, with $m \geq n$ and even $m \gg n$, which is relevant in the next-generation cellular setting.

A. Case $b \geq m + n$

As mentioned above, when $b \geq m + n$, the optimal power allocation over the transmitter antennas is given by a diagonal matrix, $\mathbf{D}$, with entries almost surely equal to 1. Under these assumptions, the following results hold.

Proposition 6: Consider a channel as in (23), affected by i.i.d. block-Rayleigh fading and with input given by (44). Let $\mathbf{\Delta} = \gamma c\mathbf{D}(\mathbf{I} + \gamma c\mathbf{D})^{-1} = \mathrm{diag}(\delta_1, \ldots, \delta_n)$, with δ_i's being distinct values. Then,

- for $m \leq n$, the pdf of its matrix-variate output, conditioned on $\mathbf{D}$ and for $n \leq b$, can be expressed as

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_m(b)K(\mathbf{Y})|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-n}|\mathbf{G}|}{\pi_m(b-n)!^m\mathcal{V}(\mathbf{\Delta})|\mathbf{I}+\gamma c\mathbf{D}|^m} \tag{46}$$

  where for $j = 1, \ldots, n$

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-n+1; y_i\delta_j) & i = 1, \ldots, m \\ \delta_j^{n-i} & i = m+1, \ldots, n \end{cases}$$

- for $m > n$, the conditioned output pdf becomes

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_n(b)K(\mathbf{Y})|\mathbf{\Delta}|^{n-m}|\mathbf{G}|}{\pi_n(b-m)!^n\mathcal{V}(\mathbf{\Delta})|\mathbf{I}+\gamma c\mathbf{D}|^m} \tag{47}$$

  where $i{=}1, \ldots, m$, are given by

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-m+1; y_i\delta_j) & j = 1, \ldots, n \\ y_i^{m-j} & j = n+1, \ldots, m\,. \end{cases}$$

  Proof: The proof is given in Appendix H. ■

1) Case $\mathbf{D} = \mathbf{I}$*:* The expressions of $p(\mathbf{Y}|\mathbf{D})$ in (46) and (47) hold provided that the diagonal elements of $\mathbf{D}$ are distinct. Thus, in general, the unconditional pdf of $\mathbf{Y}$ can be derived by integrating $p(\mathbf{Y}|\mathbf{D})$ over the distribution of $\mathbf{D}$. In this section, however, we focus on a particular power allocation matrix, $\mathbf{D} = \mathbf{I}$, and, by (45), we consider $c = b$. Note that, in this case the elements of $\mathbf{D}$ are not distinct, and expressions (46) and (47) cannot be directly evaluated. Indeed, $|\mathbf{G}| = 0$ and $\mathcal{V}(\mathbf{\Delta}) = 0$, and again a limit procedure must be applied.

We first observe that, for $\mathbf{D} = \mathbf{I}$ and $c = b$, we have $\mathbf{\Delta} = \gamma b \mathbf{D}(\mathbf{I} + \gamma b \mathbf{D})^{-1} = \bar{\delta}\mathbf{I}$ where $\bar{\delta} = \frac{\gamma b}{1+\gamma b}$.

- For $m \leq n$, we apply the limit in (4) to the ratio $|\mathbf{G}|/\mathcal{V}(\mathbf{\Delta})$ in (46) and, after some algebra, obtain

$$\lim_{\mathbf{\Delta}\to\bar{\delta}\mathbf{I}} \frac{|\mathbf{G}|}{\mathcal{V}(\mathbf{\Delta})} = \frac{\pi_n \Gamma_m(n)(b-n)!^m}{\Gamma_n(n)\Gamma_m(b)} |\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{n-m} |\widehat{\mathbf{G}}|$$

  where $\widehat{\mathbf{G}}$ is an $m \times m$ matrix whose elements are given by $(\widehat{\mathbf{G}})_{ij} = y_i^{m-j}\, {}_1F_1(n-j+1; b-j+1; y_i\bar{\delta})$, $i = 1, \ldots, m$, $j = 1, \ldots, m$. By recalling (46), the distribution of $\mathbf{Y}$ is then given by

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_m(n)}{\pi_m \Gamma_n(n)} \frac{K(\mathbf{Y})|\widehat{\mathbf{G}}|}{(1+\gamma b)^{nm}} . \quad (48)$$

- For $m > n$, we apply the limit in (4) to (47) and obtain

$$\lim_{\mathbf{\Delta}\to\bar{\delta}\mathbf{I}} \frac{|\mathbf{G}|}{\mathcal{V}(\mathbf{\Delta})} = \frac{\pi_n (b-m)!^n |\widehat{\mathbf{G}}|}{\Gamma_n(b-m+n)}$$

  where in this case

$$(\widehat{\mathbf{G}})_{ij} = y_i^{n-j}\, {}_1F_1(n-j+1; b-m+n-j+1; y_i\bar{\delta})$$

  for $i = 1, \ldots, m, j = 1, \ldots, n$, and $(\widehat{\mathbf{G}})_{ij} = y_i^{m-j}$ for $i = 1, \ldots, m, j = n+1, \ldots, m$.
  By recalling (47), it follows that

$$p(\mathbf{Y}) = \frac{\Gamma_n(b) K(\mathbf{Y}) \bar{\delta}^{n(n-m)} |\widehat{\mathbf{G}}|}{\Gamma_n(b-m+n)(1+\gamma b)^{nm}} . \quad (49)$$

We remark that, under the above assumptions, the output pdf also appears in [4]. The corresponding derivations provided therein involve Fourier integrals and Hankel matrices thus resulting in a slightly less compact form than ours.

B. A massive MIMO *regime:* $b < m + n$

Now, we consider the case of $b < m + n$; an instance of this scenario, by letting $m \gg n$, can adequately model the reverse link of the celebrated massive-MIMO channel [24]. In presence of uncorrelated block-Rayleigh fading, the high-SNR capacity-achieving input structure, as already mentioned, departs from the equal power allocation and is Beta distributed. We provide herein the output pdf for a block-fading channel fed by BSTM [5].

Proposition 7: Given a channel as in (23), with $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$, $n \leq b$, and $\mathbf{D} \sim \mathcal{B}_n(b-n, n+m-b)$, the pdf of its output can be written as

$$p(\mathbf{Y}) = \frac{\pi_n \Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)} K(\mathbf{Y}) |\mathbf{F}_4||\mathbf{Z}|}{\gamma_n c^{n(n-1)/2} \Gamma_n(n)\Gamma_n(b-n)\Gamma_n(n+m-b)} \quad (50)$$

where $\mathbf{Z}$ is an $n \times n$ matrix, whose generic entry is given by:

$$(\mathbf{Z})_{ij} = \int_0^1 \frac{(1-x)^{m-b} x^{i-1-n}}{(1+c\gamma x)^{m-b+1}} \cdot \left[\mathrm{e}^{y_j \frac{c\gamma x}{1+c\gamma x}} - \sum_{\ell,k=1}^{b-n} \frac{(\mathbf{F}_4^{-1})_{\ell k}}{y_j^{n+k-b}} \mathrm{e}^{y_{\ell+n} \frac{c\gamma x}{1+c\gamma x}} \right] \mathrm{d}x \quad (51)$$

with $(\mathbf{F}_4)_{ij} = y_{n+i}^{b-n-j}$ $i, j = 1, \ldots, b-n$.

Proof: The proof is given in Appendix I. ∎

C. Exploitation of the analytical results

We now use the above results to compute the achievable mutual information in a massive MIMO case. In order to derive the output differential entropy conditioned on the input signal, $h(\mathbf{Y}|\mathbf{X})$, we exploit the analytic expression of the conditional pdf of the output, $p(\mathbf{Y}|\mathbf{X})$, obtained above.

Proposition 8: Given a channel as in (23), the differential entropy of the output, $\mathbf{Y}$, conditioned on the channel input, $\mathbf{X}$, can be written as:

$$h(\mathbf{Y}|\mathbf{X}) = bm\log_2(\pi \mathrm{e}) + Km \sum_{i,j=1}^{n} a_{ij} \sum_{\ell=0}^{m-b} \frac{(-1)^\ell \binom{m-b}{\ell}}{s_{i,j,\ell}-1} \cdot \left[\log_2(1+c\gamma) - \frac{c\gamma\, {}_2F_1(1, s_{i,j,\ell}; s_{i,j,\ell}+1; -\gamma)}{s_{i,j,\ell} \ln 2} \right] \quad (52)$$

where K is a constant term, $s_{i,j,\ell} = b-2n+i+j+\ell$, and a_{ij} is the (i,j)-cofactor of an $n \times n$ matrix $\mathbf{A}$ such that

$$\mathbf{A}_{\ell k} = \frac{\Gamma(b-2n+\ell+k-1)\Gamma(m-n+1)}{\Gamma(b-3n+m+\ell+k)} .$$

Proof: The proof is given in Appendix J. ∎

The mutual information obtained in a massive-MIMO-like case is shown in the following figures. Fig. 4 depicts the mutual information for $n = 1$, as the SNR varies and m grows up to very large values. The plot also compares our results (denoted by markers) are compared to the approximation given in [5] for the high SNR regime (dashed lines). The two sets of curves match very closely for any value of the parameters, as expected due to the tightness of [5, eq. (8)]. As m varies, all three curves have the same slope, as this has been proven to be insensitive to the number of receiving antennas in our setting [5, eq. (8)]. As expected, better performance is obtained as m increases. However, interestingly, Fig. 5 shows that a much higher improvement can be achieved as the fading coherence length and the number of antennas at the transmitter sightly increase while m is fixed to 10. In particular, by comparing the two plots, a limited gain in performance is obtained when m increases, while, as expected, the mutual information growth is significant when n is increased by 1.

V. Conclusion

We obtained new, closed-form expressions for the probability density function of the output signal of a block-fading MIMO channel. By relying on recent results from the field of finite-dimensional random matrix theory, we provided results for the case of an i.i.d. Gaussian input under the

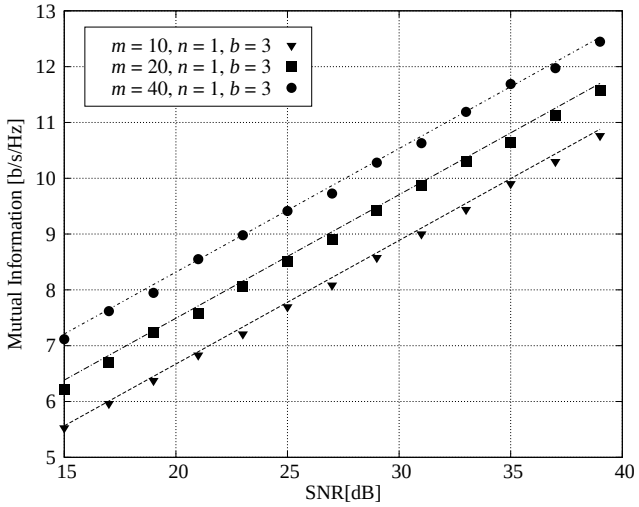

Fig. 4. Mutual information vs. SNR in massive MIMO channel with BSTM: $b = 3$, $n = 1$ and different values of m. Our results (denoted by markers) are compared to the approximation in [5] (dashed lines).

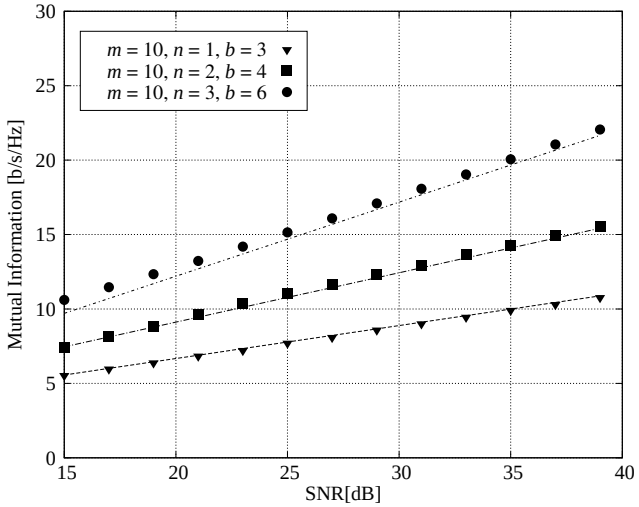

Fig. 5. Mutual information vs. SNR in massive MIMO channel with BSTM: $m = 10$ and different values of b and n. Our results (denoted by markers) are compared to the approximation in [5] (dashed lines).

assumption that the Gramian of the channel matrix is unitarily invariant. We addressed both the cases of Rayleigh and Rician fading. Furthermore, we derived the output probability density function in the case of product-form input. We particularized our newly derived expressions to those already available in the literature for the canonical case of uncorrelated Rayleigh fading, and we characterized the output signal behavior under different assumptions on the amplitude fading distribution.

VI. Acknowledgments

This paper was made possible by NPRP grant ♯5-782-2-322 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Appendix A
Proof of Lemma 1

Let $\mathbf{H}_1$ and $\mathbf{H}_2$ be, respectively, an $m \times n$ and an $m \times p$ ($m \leq p$) Gaussian complex random matrix whose columns are independent, have zero mean, and covariance $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, respectively.

- For $m \leq n$, the distribution of the ordered eigenvalues of $(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1/2}$ is given by [11, eq. (98)]

$$p(\mathbf{\Lambda}) = \frac{\pi_m^2 \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}|\{{}_1F_0(p+n-m+1;\,;-\frac{\lambda_j}{\omega_i})\}|}{(p+n-m)!^{-m}\Gamma_m(p)\Gamma_m(n)|\mathbf{\Omega}|^n\mathcal{V}(-\mathbf{\Omega}^{-1})} \tag{53}$$

where $\mathbf{\Omega} = \mathbf{\Theta}_1\mathbf{\Theta}_2^{-1}$, and $\omega_1, \ldots, \omega_m$ are the eigenvalues of $\mathbf{\Omega}$. When $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are scalar matrices, and $\mathbf{\Omega} = \mathbf{\Theta}_1\mathbf{\Theta}_2^{-1} = \omega\mathbf{I}$, the distribution of $\mathbf{\Lambda}$ can be obtained first by applying the limit (4) to (53):

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_m^2(p+n-m)!^m}{\Gamma_m(p)\Gamma_m(n)}\frac{\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\omega^{mn}} \\ &\quad \cdot \lim_{\mathbf{\Omega}\to\omega\mathbf{I}} \frac{|\{{}_1F_0(p+n-m+1;\,;-\lambda_j/\omega_i)\}|}{\mathcal{V}(-\mathbf{\Omega}^{-1})} \\ &= \frac{\pi_m^2\Gamma_m(p+n)\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\Gamma_m(m)\Gamma_m(p)\Gamma_m(n)\omega^{mn}} \\ &\quad \cdot|\{\lambda_j^{m-i}{}_1F_0(p+n-j+1;\,;-\lambda_j/\omega)\}| \end{aligned} \tag{54}$$

and then by observing that

$$\begin{aligned} &|\{\lambda_j^{m-i}{}_1F_0(p+n-i+1;\,;-\lambda_j/\omega)\}| \\ &= |\{\lambda_j^{m-i}(1+\lambda_j/\omega)^{-(p+n-i+1)}\}| \\ &= |\{[\lambda_j/(1+\lambda_j/\omega)]^{m-i}(1+\lambda_j/\omega)^{-(p+n-m+1)}\}| \\ &= \mathcal{V}(\mathbf{\Lambda}(\mathbf{I}+\mathbf{\Lambda}/\omega))|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n-m+1)} \\ &= \mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}/\omega|^{1-m}|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n-m+1)} \\ &= \mathcal{V}(\mathbf{\Lambda})|\mathbf{I}+\mathbf{\Lambda}/\omega|^{-(p+n)}\,. \end{aligned} \tag{55}$$

- For $m > n$, the distribution of the $n \times n$ random matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is given by [25, eq. (61)]

$$p(\mathbf{W}) = \frac{\Gamma_m(n+p){}_1F_0(p+n;\,;\mathbf{I}-\mathbf{\Omega}^{-1},\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1})}{\Gamma_m(p)\Gamma_n(m)|\mathbf{\Omega}|^n|\mathbf{W}|^{n-m}|\mathbf{I}+\mathbf{W}|^{p+n}} \tag{56}$$

where $\mathbf{\Omega} = \mathbf{\Theta}_1^{1/2}\mathbf{\Theta}_2^{-1}\mathbf{\Theta}_1^{1/2}$ is of size $m \times m$. Note that, for any unitary matrix $\mathbf{V}$ independent of $\mathbf{W}$, we have $|\mathbf{V}\mathbf{W}\mathbf{V}^{\mathsf{H}}| = |\mathbf{W}|$, $|\mathbf{I}+\mathbf{V}\mathbf{W}\mathbf{V}^{\mathsf{H}}| = |\mathbf{I}+\mathbf{W}|$. Moreover, the eigenvalues of $\mathbf{V}\mathbf{W}\mathbf{V}^{\mathsf{H}}(\mathbf{I}+\mathbf{V}\mathbf{W}\mathbf{V}^{\mathsf{H}})^{-1}$ are the same as those of $\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1}$. Thus, $p(\mathbf{V}\mathbf{W}\mathbf{V}) = p(\mathbf{W})$. It follows that $\mathbf{W}$ is unitarily invariant.
Let $\mathbf{\Psi} = \mathbf{W}(\mathbf{I}+\mathbf{W})^{-1}$ and the eigenvalues of $\mathbf{W}$ be $\lambda_1, \ldots, \lambda_n$. Then, $\mathbf{\Psi}$ has eigenvalues $\psi_j = \lambda_j/(1+\lambda_j)$, for $j = 1, \ldots, n$. In order to compute the hypergeometric function of two matrix arguments of different size appearing in (56), we extend $\mathbf{\Psi}$ to the $m \times m$ matrix $\widetilde{\mathbf{\Psi}}$ given by

$$\widetilde{\mathbf{\Psi}} = \begin{bmatrix} \mathbf{\Psi} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}$$

where $\mathbf{E}$ is an $(m-n)\times(m-n)$ matrix whose eigenvalues are $\mathbf{e} = [e_1, \ldots, e_{m-n}]^{\mathsf{T}}$. Then the eigenvalues of $\widetilde{\mathbf{\Psi}}$ are

$\tilde{\boldsymbol{\psi}} = [\psi_1, \ldots, \psi_n, e_1, \ldots, e_{m-n}]^{\mathsf{T}}$. It follows that

$$\begin{aligned} & {}_1F_0(p+n;;\mathbf{I}-\boldsymbol{\Omega}^{-1},\boldsymbol{\Psi}) \\ &= \lim_{\mathbf{e}\to\mathbf{0}} {}_1F_0(p+n;;\mathbf{I}-\boldsymbol{\Omega}^{-1},\widetilde{\boldsymbol{\Psi}}) \\ &\overset{(a)}{=} \lim_{\mathbf{e}\to\mathbf{0}} \frac{\Gamma_m(m)(p+n-m)!^m}{\Gamma_m(p+n)} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\mathbf{I}-\boldsymbol{\Omega}^{-1})\mathcal{V}(\widetilde{\boldsymbol{\Psi}})} \\ &= \frac{\Gamma_m(m)(p+n-m)!^m}{\Gamma_m(p+n)\mathcal{V}(\mathbf{I}-\boldsymbol{\Omega}^{-1})} \lim_{\mathbf{e}\to\mathbf{0}} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\widetilde{\boldsymbol{\Psi}})} \end{aligned} \quad (57)$$

where in (a) we applied (6) and $f_i(\tilde{\psi}_j) = {}_1F_0(p+n-m+1;;(1-\omega_i^{-1})\tilde{\psi}_j)$, $i,j = 1,\ldots,m$. By applying the limit in (3) and the properties in (1) and (7), the limit in (57) can be computed as

$$\lim_{\mathbf{e}\to\mathbf{0}} \frac{|\{f_i(\tilde{\psi}_j)\}|}{\mathcal{V}(\widetilde{\boldsymbol{\Psi}})} = \frac{\Gamma_m(p+n)\Gamma_n(m)}{\Gamma_n(p+n)\Gamma_m(m)} \frac{(p+n-m)!^{n-m}|\mathbf{F}|}{|\boldsymbol{\Psi}|^{m-n}\mathcal{V}(\boldsymbol{\Psi})} \quad (58)$$

where for $i = 1,\ldots,m$

$$(\mathbf{F})_{ij} = \begin{cases} f_i(\psi_j), & j = 1,\ldots,n \\ (1-\omega_i^{-1})^{m-j} & j = n+1,\ldots,m \end{cases}$$

We now write the eigenvalue decomposition of $\mathbf{W}$ as $\mathbf{W} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{H}}$, where $(\lambda_1,\ldots,\lambda_n) = \mathrm{diag}(\boldsymbol{\Lambda})$. We then observe that $|\mathbf{W}| = |\boldsymbol{\Lambda}|$, $|\mathbf{I}+\mathbf{W}| = |\mathbf{I}+\boldsymbol{\Lambda}|$, $|\boldsymbol{\Psi}| = |\boldsymbol{\Lambda}||\mathbf{I}+\boldsymbol{\Lambda}|^{-1}$, and that $\mathcal{V}(\boldsymbol{\Psi}) = \mathcal{V}(\boldsymbol{\Lambda})|\mathbf{I}+\boldsymbol{\Lambda}|^{1-n}$. Therefore, by using (56), (57), and (58), the pdf of $\mathbf{W}$ can be rewritten as

$$p(\mathbf{W}) = \frac{\Gamma_m(p+n)(p+n-m)!^n}{\Gamma_n(p+n)\Gamma_m(p)} \frac{|\mathbf{I}+\boldsymbol{\Lambda}|^{m-p-n-1}|\mathbf{F}|}{|\boldsymbol{\Omega}|^n\mathcal{V}(\mathbf{I}-\boldsymbol{\Omega}^{-1})\mathcal{V}(\boldsymbol{\Lambda})}\,. \quad (59)$$

The pdf of the ordered eigenvalues of a complex random $n\times n$ matrix $\mathbf{W}$ is given by [11, eq. (93)]:

$$p(\boldsymbol{\Lambda}) = \frac{\pi_n^2\mathcal{V}(\boldsymbol{\Lambda})^2}{\Gamma_n(n)} \int p_{\mathbf{W}}(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{H}})\,\mathrm{d}\mathbf{U}\,.$$

In our case, since $p(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{H}})$ does not depend on $\mathbf{U}$, we obtain (14).

Appendix B
Proof of Lemma 2

- For $m \le n$, let us define $\mathbf{W}_2 = \mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}}$. Then the matrix $\mathbf{W} = \mathbf{W}_2^{-1/2}\mathbf{H}_1\mathbf{H}_1{}^{\mathsf{H}}\mathbf{W}_2^{-1/2}$ can be rewritten as $\mathbf{W} = \mathbf{H}\mathbf{H}^{\mathsf{H}}$ where $\mathbf{H} = \mathbf{W}_2^{-1/2}\mathbf{H}_1$. For any given matrix $\mathbf{W}_2$, for $\boldsymbol{\Theta} = \theta\mathbf{I}$ and $\mathbf{M}\mathbf{M}^{\mathsf{H}} = \mu\mathbf{I}$, $\mathbf{H}$ is a Gaussian complex matrix with average $\widetilde{\mathbf{M}} = \mathbf{W}_2^{-1/2}\mathbf{M}$ and independent columns whose covariance is $\boldsymbol{\Sigma} = \theta\mathbf{W}_2^{-1}$. It follows that, given $\mathbf{W}_2$, $\mathbf{W}$ is a non central Wishart matrix and $p(\mathbf{W}|\mathbf{W}_2)$ is given by [11, eq. (99)]

$$\begin{aligned} p_{\mathbf{W}|\mathbf{W}_2}(\mathbf{W}|\mathbf{W}_2) &= {}_0F_1(;n;\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\mathbf{W}) \\ &\quad \cdot\frac{\mathrm{e}^{-\mathsf{Tr}\{\boldsymbol{\Sigma}^{-1}\mathbf{W}\}}|\mathbf{W}|^{n-m}}{\mathrm{e}^{\mathsf{Tr}\{\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^{\mathsf{H}}\}}\Gamma_m(n)|\boldsymbol{\Sigma}|^n} \\ &= {}_0F_1(;n;\theta^{-2}\mu\mathbf{W}_2\mathbf{W}) \\ &\quad \cdot\frac{\mathrm{e}^{-\mathsf{Tr}\{\theta^{-1}\mathbf{W}_2\mathbf{W}\}}|\mathbf{W}_2|^n}{\mathrm{e}^{\mu m/\theta}\theta^{nm}\Gamma_m(n)|\mathbf{W}|^{m-n}}\,. \end{aligned}$$

On the other hand, $\mathbf{W}_2$ is a central Wishart with covariance $\theta\mathbf{I}$. Thus, the density of $\mathbf{W}$ can be written as

$$\begin{aligned} p_{\mathbf{W}}(\mathbf{W}) &= \int p_{\mathbf{W}|\mathbf{W}_2}(\mathbf{W}|\mathbf{W}_2)p_{\mathbf{W}_2}(\mathbf{W}_2)\,\mathrm{d}\mathbf{W}_2 \\ &= \int {}_0F_1(;n;\frac{\mu}{\theta^2}\mathbf{W}_2\mathbf{W}) \\ &\quad \cdot\frac{\mathrm{e}^{-\mathsf{Tr}\{\mathbf{W}_2(\mathbf{I}+\mathbf{W})/\theta\}}|\mathbf{W}_2|^{p+n-m}\,\mathrm{d}\mathbf{W}_2}{\mathrm{e}^{\mu m/\theta}\theta^{(p+n)m}\Gamma_m(n)\Gamma_m(p)|\mathbf{W}|^{m-n}} \\ &= \frac{|\mathbf{W}|^{n-m}}{\mathrm{e}^{\mu m/\theta}\theta^{(p+n)m}\Gamma_m(n)\Gamma_m(p)} \\ &\quad \cdot\int \frac{{}_0F_1(;n;\frac{\mu}{\theta^2}\mathbf{W}_2\mathbf{W})\,\mathrm{d}\mathbf{W}_2}{\mathrm{e}^{\mathsf{Tr}\{\mathbf{W}_2(\mathbf{I}+\mathbf{W})/\theta\}}|\mathbf{W}_2|^{m-p-n}}\,. \end{aligned}$$

In order to solve the above integral, we employ the following result

$$\int_{\mathbf{B}=\mathbf{B}^{\mathsf{H}}>0} \frac{{}_pF_q(\mathbf{a};\mathbf{b};\mathbf{C}\mathbf{B})}{|\mathbf{B}|^{m-c}\mathrm{e}^{\mathsf{Tr}\{\mathbf{A}\mathbf{B}\}}}\,\mathrm{d}\mathbf{B} = \frac{{}_{p+1}F_q(\mathbf{a},c;\mathbf{b};\mathbf{C}\mathbf{A}^{-1})}{[\Gamma_m(c)]^{-1}|\mathbf{A}|^c}\,,$$

which holds for $m\times m$ matrices $\mathbf{A},\mathbf{B}$, and $\mathbf{C}$, and for $\mathbb{R}(c) > m-1$ [26, eq. (115)]. Then,

$$p_{\mathbf{W}}(\mathbf{W}) = \frac{\Gamma_m(p+n){}_1F_1(p+n;n;\frac{\mu}{\theta}\mathbf{W}(\mathbf{I}+\mathbf{W})^{-1})}{\Gamma_m(n)\Gamma_m(p)\mathrm{e}^{\mu m/\theta}|\mathbf{W}|^{m-n}|\mathbf{I}+\mathbf{W}|^{p+n}}\,.$$

It can be observed that $p(\mathbf{W})$ depends only on the eigenvalues of $\mathbf{W}$, thus it is unitarily invariant. It follows that the pdf of the ordered eigenvalues of $\mathbf{W}$ is given by [11, eq. (93)]

$$p(\boldsymbol{\Lambda}) = \frac{\pi_m^2\mathcal{V}(\boldsymbol{\Lambda})^2}{\Gamma_m(m)} \int p_{\mathbf{W}}(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{H}})\,\mathrm{d}\mathbf{U}\,,$$

which provides the results in (15).

- For $m > n$, the distribution of the $n\times n$ matrix $\mathbf{W} = \mathbf{H}_1{}^{\mathsf{H}}(\mathbf{H}_2\mathbf{H}_2{}^{\mathsf{H}})^{-1}\mathbf{H}_1$ is given by [11, eq. (105)]

$$p(\mathbf{W}) = \frac{\Gamma_n(p+n){}_1F_1(p+n;m;\boldsymbol{\Omega}(\mathbf{I}+\mathbf{W}^{-1})^{-1})}{\Gamma_n(m)\Gamma_n(p+n-m)\mathrm{e}^{\mathsf{Tr}\{\boldsymbol{\Omega}\}}|\mathbf{W}|^{n-m}|\mathbf{I}+\mathbf{W}|^{p+n}} \quad (60)$$

and the distribution of its eigenvalues is given by

$$p(\boldsymbol{\Lambda}) = \frac{p!^n}{(m-n)!^n} \frac{\pi_n\mathrm{e}^{-\mathsf{Tr}\{\boldsymbol{\Omega}\}}\mathcal{V}(\boldsymbol{\Lambda})|\mathbf{F}||\boldsymbol{\Lambda}|^{m-n}}{\Gamma_n(p+n-m)\mathcal{V}(\boldsymbol{\Omega})|\mathbf{I}+\boldsymbol{\Lambda}|^{1+p}} \quad (61)$$

where $\boldsymbol{\Omega} = \mathbf{M}^{\mathsf{H}}\boldsymbol{\Theta}^{-1}\mathbf{M}$, $(\mathbf{F})_{ij} = {}_1F_1(p+1;m-n+1;\lambda_j\omega_i/(1+\lambda_j))$, and $\omega_1,\ldots,\omega_n$ are the eigenvalues of $\boldsymbol{\Omega}$. This result has been obtained by applying (6) to [11, eq. (106)].

In the particular case where $\boldsymbol{\Omega}$ is a scalar matrix (i.e., $\boldsymbol{\Omega} = \omega\mathbf{I}$), matrix $\mathbf{W}$ is unitarily invariant since its pdf in (60) only depends on its eigenvalues $\boldsymbol{\Lambda}$. Indeed, $|\mathbf{W}| = |\boldsymbol{\Lambda}|$, $|\mathbf{I}+\mathbf{W}| = |\mathbf{I}+\boldsymbol{\Lambda}|$, and the generalized hypergeometric function ${}_1F_1(p+n;m;\omega(\mathbf{I}+\mathbf{W}^{-1})^{-1})$ only depends on the eigenvalues of its matrix argument, i.e., on $\boldsymbol{\Lambda}$. In such a case, the distribution of $\boldsymbol{\Lambda}$ can be obtained from (61) by applying the limit in (4) to the ratio $|\mathbf{F}|/\mathcal{V}(\boldsymbol{\Omega})$ and the property in (7). The result is reported in (16).

APPENDIX C
PROOF OF PROPOSITION 1

The proof of (19) and (21) follows from the application of [27, Theorem I] to (17) and (18), respectively.

The density given in (18) is an ordered eigenvalue distribution and the unordered eigenvalue distribution is obtained by dividing (18) by $n!$. Then, applying the Laplace determinant expansion, the unordered eigenvalues distribution becomes

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_n^2 \Gamma_n(p+q)(1-\lambda_1)^{q-n}\lambda_1^{p-n-2}}{n!\Gamma_n(n)\Gamma_n(p)\Gamma_n(q)} \\ &\quad \cdot \sum_{i=1}^{n}\sum_{j=1}^{n}(-\lambda_1)^{i+j} \\ &\quad \cdot \prod_{k=2}^{n}(1-\lambda_k)^{q-n}\lambda_k^{q-n}|\bar{\mathbf{V}}(\mathbf{\Lambda})||\widetilde{\mathbf{V}}(\mathbf{\Lambda})| \end{aligned} \tag{62}$$

where $\bar{\mathbf{V}}(\mathbf{\Lambda})$ and $\widetilde{\mathbf{V}}(\mathbf{\Lambda})$ are $(n-1)\times(n-1)$ matrices obtained by deleting the first row and column from the Vandermonde matrix $\mathbf{V}(\mathbf{\Lambda})$ and its conjugate transpose, separately. The (i,j)-th entry of $\mathbf{V}(\mathbf{\Lambda})$ and its conjugate transpose are λ_i^{j-1} and λ_j^{i-1}, respectively. Thanks to [28, Corollary 1], the result in (21) can be obtained through integration over $n-1$ eigenvalues from λ_2 to λ_n. The final expressions are in both cases due to the definition of the scalar Beta function [14]. It should be noticed that the choice of λ_1 in (62) has no effect on the final result, since we started from an unordered eigenvalue distribution. Using the same approach, the proof of (19) is straightforward.

APPENDIX D
PROOF OF (24) AND (26)

We first observe that for $m > n$ the matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ does not have full rank and has $m-n$ zero eigenvalues. The n non-zero eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$, denoted by $\lambda_1,\ldots,\lambda_n$, are also the eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ and are the elements of the $n\times n$ diagonal matrix $\mathbf{\Lambda}$. We start by rewriting [3, eq. (38)] in the case $m > n$ and obtain

$$p(\mathbf{Y}) = \int \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2}p(\mathbf{\Lambda})}{\pi^{mb}|\mathbf{I}+\gamma\mathbf{\Lambda}|^b}\left[\int \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{C}}\mathbf{U}^{\mathsf{H}}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\mathbf{U}\}}p(\mathbf{U}|\mathbf{\Lambda})\,\mathrm{d}\mathbf{U}\right]\mathrm{d}\mathbf{\Lambda} \tag{63}$$

where $\mathbf{U}$ is a unitary $m\times m$ matrix, $\widetilde{\mathbf{C}}$ is an $m\times m$ diagonal matrix whose elements are given by $(\widetilde{\mathbf{C}})_{jj} = c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, $j=1,\ldots,m$, with $c_j = 0$, for $j = n+1,\ldots,m$. Since we assume that $\mathbf{W} = \mathbf{H}^{\mathsf{H}}\mathbf{H}$ is unitarily invariant, its eigenvalues do not depend on $\mathbf{U}$. Moreover, $\mathbf{U}$ is a Haar matrix (see Definition 3). Then, $p(\mathbf{U}|\mathbf{\Lambda}) = p(\mathbf{U})$. The inner integral over $\mathbf{U}$ can be solved using the Harish-Chandra-Itzykson-Zuber integral [29]

$$\int_{\mathcal{U}(m)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{C}}\mathbf{U}^{\mathsf{H}}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\mathbf{U}\}}p(\mathbf{U})\,\mathrm{d}\mathbf{U} = \frac{\Gamma_m(m)|\mathbf{E}|}{\pi_m\mathcal{V}(\widetilde{\mathbf{C}})\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})}\,.$$

The elements of matrix $\mathbf{E}$ are given by $(\mathbf{E})_{ij} = \mathrm{e}^{y_i c_j}$, $i,j = 1,\ldots,m$ and y_i, $i=1,\ldots,m$, are the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. Due to the fact that $c_j = 0$ for $j = n+1,\ldots,m$, we have $|\mathbf{E}| = 0$ and $\mathcal{V}(\widetilde{\mathbf{C}}) = 0$; thus the limit in (3) must be applied to the term $|\mathbf{E}|/\mathcal{V}(\widetilde{\mathbf{C}})$. We have

$$\lim_{c_{n+1},\ldots,c_m\to 0}\frac{|\mathbf{E}|}{\mathcal{V}(\widetilde{\mathbf{C}})} = \frac{\pi_m\Gamma_n(m)}{\pi_n\Gamma_m(m)}\frac{|\widetilde{\mathbf{E}}|}{\mathcal{V}(\mathbf{C})|\mathbf{C}|^{m-n}} \tag{64}$$

where $\widetilde{\mathbf{E}}$ is an $m\times m$ matrix whose elements are given by $(\widetilde{\mathbf{E}})_{ij} = \mathrm{e}^{y_i c_j}$ for $1\le j\le n$, and $(\widetilde{\mathbf{E}})_{ij} = y_i^{j-n-1}$ for $n+1\le j\le m$. Also, $\mathbf{C}$ is an $n\times n$ diagonal matrix whose elements are $(\mathbf{C})_{jj} = c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, $j=1,\ldots,n$. Therefore, (63) can be rewritten as

$$p(\mathbf{Y}) = \frac{\Gamma_n(m)K(\mathbf{Y})}{\pi_n}\int \frac{p(\mathbf{\Lambda})|\widetilde{\mathbf{E}}|}{|\mathbf{I}+\gamma\mathbf{\Lambda}|^b}\frac{|\mathbf{C}|^{n-m}}{\mathcal{V}(\mathbf{C})}\,\mathrm{d}\mathbf{\Lambda} \tag{65}$$

where $K(\mathbf{Y}) = \mathrm{e}^{-\|\mathbf{Y}\|^2}/(\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})\pi^{mb})$ was defined in (25). Since $c_j = \lambda_j\gamma/(1+\gamma\lambda_j)$, by applying the definition of the Vandermonde determinant, we get $\mathcal{V}(\mathbf{C}) = |\mathbf{I}+\gamma\mathbf{\Lambda}|^{1-n}\mathcal{V}(\gamma\mathbf{\Lambda})$. Moreover, $|\mathbf{C}| = |\gamma\mathbf{\Lambda}||\mathbf{I}+\gamma\mathbf{\Lambda}|^{-1}$. By substituting these results in (65), we obtain (26).

APPENDIX E
PROOF OF PROPOSITION 2

We first observe that the matrix $\mathbf{H}$ in (29) can be written as $\mathbf{H} = \mathbf{H}_0/\sqrt{1+\kappa}$, where $\mathbf{H}_0 = \sqrt{\kappa}\bar{\mathbf{H}} + \tilde{\mathbf{H}}$.

- For $m\le n$ and $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathsf{H}} = h\mathbf{I}$, the joint distribution of the ordered eigenvalues of $\mathbf{H}_0\mathbf{H}_0^{\mathsf{H}}$ is given by (12) where $\mu = \kappa h$, i.e.,

  $$p_0(\mathbf{\Lambda}_0) = \frac{\pi_m^2|\mathbf{\Lambda}_0|^n\mathcal{V}(\mathbf{\Lambda}_0)|\{\lambda_{0j}^{-i}{}_0F_1(\,;n-i+1;\kappa h\lambda_{0j})\}|}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\kappa hm+\mathsf{Tr}\{\mathbf{\Lambda}_0\}}}$$

  where $(\lambda_{01},\ldots,\lambda_{0m}) = \mathrm{diag}(\mathbf{\Lambda}_0)$. Then, the pdf of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is given by

  $$\begin{aligned} p(\mathbf{\Lambda}) &= (1+\kappa)^m p_0((1+\kappa)\mathbf{\Lambda}) \\ &= \frac{\pi_m^2(1+\kappa)^{mn}|\mathbf{\Lambda}|^n|\mathbf{F}|\mathcal{V}(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\kappa hm+(1+\kappa)\mathsf{Tr}\{\mathbf{\Lambda}\}}} \end{aligned} \tag{66}$$

  where $(\mathbf{F})_{ij} = \lambda_j^{-i}{}_0F_1(\,;n-i+1;\kappa(1+\kappa)h\lambda_j)$, $i,j = 1,\ldots,m$. By substituting this equation in (24) and by applying the result in Appendix K, we obtain (30).
- For $m > n$, and for $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}} = h\mathbf{I}$, we adopt a procedure similar to the one above. In this case, the pdf of the non-zero eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is given by (66) where n and m should be replaced by m and n, respectively. By substituting $p(\mathbf{\Lambda})$ in (26) and by applying the result in Appendix K, we obtain (31).

APPENDIX F
PROOF OF PROPOSITION 3

For $m\le n$, the distribution of the ordered eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is expressed as [9, eq. (9)]

$$p(\mathbf{\Lambda}) = \frac{\Gamma(\alpha-m+1)^m}{\Gamma(n-m+1)^m}\frac{\pi_m\mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}}{\mathcal{V}((\mathbf{I}+\mathbf{\Omega})^{-1})}\frac{\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}|\mathbf{F}|}{\Gamma_m(\alpha)|\mathbf{I}+\mathbf{\Omega}^{-1}|^\alpha} \tag{67}$$

with $(\mathbf{F})_{ij} = {}_1F_1(\alpha-m+1;n-m+1;\lambda_j/(1+\omega_i))$. When $\mathbf{\Omega} = \omega\mathbf{I}$, the expression of $p(\mathbf{\Lambda})$ can be derived from (67) by applying the limit in (4) and by using the property in (1).

For simplicity, we define $\mathbf{\Theta} = (\mathbf{I} + \mathbf{\Omega})^{-1} = \theta\mathbf{I}$ where $\theta = (1+\omega)^{-1}$. Then,

$$\begin{aligned} p(\mathbf{\Lambda}) &= \frac{\pi_m \Gamma(\alpha-m+1)^m \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^{n-m}}{\Gamma_m(\alpha)\Gamma(n-m+1)^m \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}\}}(1+1/\omega)^{m\alpha}} \\ &\quad \cdot \lim_{\mathbf{\Theta}\to\theta\mathbf{I}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Theta})} \\ &= \frac{\pi_m^2 \mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^n |\widetilde{\mathbf{F}}|}{\Gamma_m(m)\Gamma_m(n)\mathrm{e}^{\mathsf{Tr}\{\mathbf{\Lambda}\}}(1+1/\omega)^{m\alpha}} \end{aligned} \tag{68}$$

where $(\widetilde{\mathbf{F}})_{ij} = \lambda_j^{-i}{}_1F_1(\alpha-i+1; n-i+1; \lambda_j/(1+\omega))$, $i,j = 1,\ldots,m$. The proposition statement follows by replacing (68) in (24). Similarly, when $m > n$ and $\mathbf{\Omega} = \omega\mathbf{I}$, the distribution of the eigenvalues of $\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is given by

$$p(\mathbf{\Lambda}) = \frac{\pi_n^2 \mathrm{e}^{-\mathsf{Tr}\{\mathbf{\Lambda}\}}\mathcal{V}(\mathbf{\Lambda})|\mathbf{\Lambda}|^m}{\Gamma_n(m)(1+1/\omega)^{n\alpha}} \frac{|\widetilde{\mathbf{F}}|}{\Gamma_n(n)}$$

where $(\widetilde{\mathbf{F}})_{ij} = \lambda_j^{-i}{}_1F_1(\alpha-i+1; m-i+1; \lambda_j/(1+\omega))$, $i,j = 1,\ldots,n$. Again, the proposition statement is obtained by replacing the above equation in (26).

Appendix G
Proof of Proposition 5

We first observe that the matrix $\mathbf{H}$ in (35) can be written as $\mathbf{H} = \mathbf{H}_0/\sqrt{1+\kappa}$, where

$$\mathbf{H}_0 = \sqrt{\kappa}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\bar{\mathbf{H}}_s + \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathsf{H}}\right)^{-1/2}\widetilde{\mathbf{H}}_s$$

- for $m \le n$, and $\bar{\mathbf{H}}_s\bar{\mathbf{H}}_s{}^{\mathsf{H}} = h\mathbf{I}$, the distribution of the ordered eigenvalues of $\mathbf{H}_0\mathbf{H}_0{}^{\mathsf{H}}$ is given by (15)

$$\begin{aligned} p_0(\mathbf{\Lambda}_0) &= \frac{\pi_m^2 \mathrm{e}^{-h\kappa m/\theta}\Gamma_m(Ln+n)\mathcal{V}(\mathbf{\Lambda}_0)^2|\mathbf{\Lambda}_0|^{n-m}}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(Ln)|\mathbf{I}+\mathbf{\Lambda}_0|^{Ln+n}} \\ &\quad \cdot {}_1F_1\left(Ln+n; n; \frac{h\kappa}{\theta}\mathbf{\Lambda}_0(\mathbf{I}+\mathbf{\Lambda}_0)^{-1}\right) \end{aligned}$$

where we set $\mu = h\kappa$. The distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can then be obtained as

$$\begin{aligned} p(\mathbf{\Lambda}) &= \tilde{\kappa}^m p_0(\tilde{\kappa}\mathbf{\Lambda}) \\ &= \frac{\pi_m^2 \tilde{\kappa}^{mn}\Gamma_m(Ln+n)|\mathbf{\Lambda}|^{n-m}\mathcal{V}^2(\mathbf{\Lambda})}{\Gamma_m(m)\Gamma_m(n)\Gamma_m(Ln)\mathrm{e}^{h\kappa m/\theta}|\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda}|^{Ln+n}} \\ &\quad \cdot {}_1F_1\left(Ln+n; n; \frac{h\kappa\tilde{\kappa}}{\theta}\mathbf{\Lambda}(\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda})^{-1}\right) \end{aligned} \tag{69}$$

where $\tilde{\kappa} = 1+\kappa$. By substituting the above expression in (24), and by exploiting the property [12, eq. (2.36)]

$${}_1F_1(a; b; \mathbf{\Psi}) = \frac{|\{{}_1F_1(a-m+j; b-m+j; \psi_i)\psi_i^{j-1}\}|}{\mathcal{V}(\mathbf{\Psi})},$$

which holds for any $m \times m$ Hermitian matrix $\mathbf{\Psi}$ with eigenvalues $\psi_1,\ldots,\psi_m$, we obtain (39).

- For $m > n$, the distribution of the eigenvalues of $\mathbf{H}_0\mathbf{H}_0{}^{\mathsf{H}}$ is given by (16):

$$p_0(\mathbf{\Lambda}_0) = \frac{\pi_n^2\Gamma_n(Ln+n)\mathrm{e}^{-\omega n}|\mathbf{F}||\mathbf{\Lambda}_0|^{m-n}\mathcal{V}(\mathbf{\Lambda}_0)}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(Ln+n-m)|\mathbf{I}+\mathbf{\Lambda}_0|^{Ln+1}}$$

where $\mathbf{\Omega} = \omega\mathbf{I} = \mathbf{M}^{\mathsf{H}}\mathbf{\Theta}^{-1}\mathbf{M}$. In our case we have $\mathbf{M} = \sqrt{\kappa}\bar{\mathbf{H}}_s$, thus $\omega\mathbf{I} = \kappa\bar{\mathbf{H}}_s{}^{\mathsf{H}}\mathbf{\Theta}^{-1}\bar{\mathbf{H}}_s$. It follows that the matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is unitarily invariant if $\bar{\mathbf{H}}_s{}^{\mathsf{H}}\mathbf{\Theta}^{-1}\bar{\mathbf{H}}_s = \omega/\kappa\mathbf{I}$. The distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ can then be obtained as

$$\begin{aligned} p(\mathbf{\Lambda}) &= \tilde{\kappa}^n p_0(\tilde{\kappa}\mathbf{\Lambda}) \\ &= \frac{\pi_n^2\Gamma_n(Ln+n)|\mathbf{\Lambda}|^{m-n}\mathcal{V}(\mathbf{\Lambda})}{\Gamma_n(n)\Gamma_n(m)\Gamma_n(Ln+n-m)} \\ &\quad \cdot \frac{|\{\tilde{\lambda}_j^{n-i}{}_1F_1(Ln+n-i+1; m-i+1; \omega\tilde{\kappa}\tilde{\lambda}_j)\}|}{\tilde{\kappa}^{-nm}\mathrm{e}^{\omega n}|\mathbf{I}+\tilde{\kappa}\mathbf{\Lambda}|^{Ln+1}} \end{aligned}$$

where $\tilde{\kappa} = 1+\kappa$ and $\tilde{\lambda}_j = \lambda_j/(1+\tilde{\kappa}\lambda_j)$. By substituting this expression in (26), we obtain (41).

Appendix H
Proof of Proposition 6

Given the above assumptions and considering that $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$, $p(\mathbf{Y}|\mathbf{D})$ is given in as [5, eq. (53)]:

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\mathrm{e}^{-\|\mathbf{Y}\|^2}A}{\pi^{mb}|\mathbf{I}+\gamma c\mathbf{D}|^m} \tag{70}$$

where

$$\begin{aligned} A &= \int_{\mathcal{S}(b,n)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}} p(\mathbf{\Phi})\,\mathrm{d}\mathbf{\Phi} \\ &= \frac{1}{|\mathcal{S}(b,n)|}\int_{\mathcal{S}(b,n)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}}\,\mathrm{d}\mathbf{\Phi} \end{aligned}$$

and $\mathbf{\Delta} = \gamma c\mathbf{D}(\mathbf{I}+\gamma c\mathbf{D})^{-1}$. In [5, Appendix A], it is observed that the integral above is not an instance of the Harish-Chandra-Itzykson-Zuber (HCIZ) integral [29] since the $n \times b$ matrix $\mathbf{\Phi}$ is not a square matrix. In order to circumvent this problem, one has to extend matrix $\mathbf{\Phi}^{\mathsf{H}}$ to the unitary $b \times b$ Haar matrix $\widetilde{\mathbf{\Phi}}^{\mathsf{H}} = [\mathbf{\Phi}^{\mathsf{H}}, \mathbf{\Phi}_\perp{}^{\mathsf{H}}]$, where $\mathbf{\Phi}_\perp{}^{\mathsf{H}}$ is the orthogonal complement of $\mathbf{\Phi}^{\mathsf{H}}$ with respect to the unitary group $\mathcal{U}(b)$. Thus, following [5, Appendix A], we can write

$$A = \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|}\int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\mathbf{\Delta}\mathbf{\Phi}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\mathbf{\Phi}^{\mathsf{H}}\}}\,\mathrm{d}\widetilde{\mathbf{\Phi}}\,.$$

Next, the $n \times n$ diagonal matrix $\mathbf{\Delta} = \mathrm{diag}(\delta_1,\ldots,\delta_n)$ can be extended to the $b \times b$ matrix $\widetilde{\mathbf{\Delta}} = \mathrm{diag}(\delta_1,\ldots,\delta_n,q_1,\ldots,q_{b-n})$ where the elements of $\mathbf{q} = [q_1,\ldots,q_{b-n}]$ are distinct and different from $\delta_1,\ldots,\delta_n$. The above integral can then be written as

$$A = \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|}\lim_{\mathbf{q}\to 0}\int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{\Phi}}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}\widetilde{\mathbf{\Phi}}^{\mathsf{H}}\}}\,\mathrm{d}\widetilde{\mathbf{\Phi}}\,. \tag{71}$$

We observe that the matrix $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$ has $b-m$ zero-eigenvalues, and its non-zero eigenvalues are the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. Since the HCIZ integral is a function of the eigenvalues of the matrices $\widetilde{\mathbf{\Delta}}$ and $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$, we replace the matrix $\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$ with the $b \times b$ block diagonal matrix $\mathbf{\Psi} = \mathrm{diag}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}}, \mathbf{P})$ where $\mathbf{P}$ is diagonal and has diagonal entries $\mathbf{p} = [p_1,\ldots,p_{b-m}]$. Such

elements are positive, distinct, and they are different from the eigenvalues of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$. In conclusion, we can write:

$$\begin{aligned} A &= \frac{1}{|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \lim_{\mathbf{q}\to\mathbf{0}} \lim_{\mathbf{p}\to\mathbf{0}} \int_{\mathcal{U}(b)} \mathrm{e}^{\mathsf{Tr}\{\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{\Phi}}\mathbf{\Psi}\widetilde{\mathbf{\Phi}}^{\mathsf{H}}\}}\, \mathrm{d}\widetilde{\mathbf{\Phi}} \\ &= \frac{\Gamma_b(b)|\mathcal{U}(b)|}{\pi_b|\mathcal{S}(b,n)||\mathcal{U}(b-n)|} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} \\ &= \frac{\Gamma_b(b)}{\pi_b} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} \end{aligned} \quad (72)$$

where $(\mathbf{F})_{ij} = \mathrm{e}^{\psi_i\tilde{\delta}_j}$ and ψ_i and $\tilde{\delta}_j$ are the eigenvalues of $\mathbf{\Psi}$ and $\widetilde{\mathbf{\Delta}}$, respectively. In (72) we first used the HCIZ integral [29] and then the equality $|\mathcal{U}(b)| = |\mathcal{S}(b,n)||\mathcal{U}(b-n)|$. Then, we apply twice the limit in (3) and obtain:

$$\begin{aligned} \lim_{\mathbf{q}\to 0} \lim_{\mathbf{p}\to\mathbf{0}} \frac{|\mathbf{F}|}{\mathcal{V}(\mathbf{\Psi})\mathcal{V}(\widetilde{\mathbf{\Delta}})} &= \frac{\pi_b\Gamma_m(b)|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-b}}{\pi_m\Gamma_b(b)\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}})} \\ &\quad \cdot \frac{\pi_b\Gamma_n(b)|\widehat{\mathbf{F}}||\mathbf{\Delta}|^{n-b}}{\pi_n\Gamma_b(b)\mathcal{V}(\mathbf{\Delta})} \end{aligned} \quad (73)$$

where for $m \le n$,

$$(\widehat{\mathbf{F}})_{ij} = \begin{cases} \mathrm{e}^{y_i\delta_j} & i=1,\ldots,m; j=1,\ldots,n \\ y_i^{b-j} & i=1,\ldots,m; j=n+1,\ldots,b \\ \delta_j^{b-i} & i=m+1,\ldots,b, j=1,\ldots,n \\ (b-i)! & i=j; j=n+1,\ldots,b \\ 0 & \text{elsewhere}, \end{cases} \quad (74)$$

while for $m > n$,

$$(\widehat{\mathbf{F}})_{ij} = \begin{cases} \mathrm{e}^{y_i\delta_j} & i=1,\ldots,m; j=1,\ldots,n \\ y_i^{b-j} & i=1,\ldots,m; j=n+1,\ldots,b \\ \delta_j^{b-i} & i=m+1,\ldots,b, j=1,\ldots,n \\ (b-i)! & i=j; j=m+1,\ldots,b \\ 0 & \text{elsewhere.} \end{cases} \quad (75)$$

In summary,

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\pi_b\Gamma_m(b)\Gamma_n(b)K(\mathbf{Y})|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{m-b}|\widehat{\mathbf{F}}||\mathbf{\Delta}|^{n-b}}{\pi_m\pi_n\Gamma_b(b)\mathcal{V}(\mathbf{\Delta})|\mathbf{I}+\gamma c\mathbf{D}|^m} \quad (76)$$

where $K(\mathbf{Y})$ was defined in (25).

We now focus on the case $m > n$ and compute the determinant $|\widehat{\mathbf{F}}|$. Note that $\widehat{\mathbf{F}}$ can be written as

$$\widehat{\mathbf{F}} = \begin{bmatrix} \widehat{\mathbf{F}}_1 & \widehat{\mathbf{F}}_2 \\ \widehat{\mathbf{F}}_3 & \widehat{\mathbf{F}}_4 \end{bmatrix}$$

where $\widehat{\mathbf{F}}_1$ is of size $m\times m$, $\widehat{\mathbf{F}}_2$ $m\times(b-m)$, $\widehat{\mathbf{F}}_3$ $(b-m)\times m$, and $\widehat{\mathbf{F}}_4$ $(b-m)\times(b-m)$. By using the property of the determinant of block matrices [30], we have:

$$|\widehat{\mathbf{F}}| = |\widehat{\mathbf{F}}_4||\widehat{\mathbf{T}}|\,,$$

where $\widehat{\mathbf{T}} = \widehat{\mathbf{F}}_1 - \widehat{\mathbf{F}}_2\widehat{\mathbf{F}}_4^{-1}\widehat{\mathbf{F}}_3$. In our case, $\widehat{\mathbf{F}}_4$ is diagonal (see the definition of $\widehat{\mathbf{F}}$ in (75)) and $|\widehat{\mathbf{F}}_4| = \prod_{i=0}^{b-m-1} i!$. Moreover, we have $(\widehat{\mathbf{F}}_2\widehat{\mathbf{F}}_4^{-1}\widehat{\mathbf{F}}_3)_{ij} = \sum_{k=0}^{b-m-1}(y_i\delta_j)^k/k!$ for $i=1,\ldots,m$, $j=1,\ldots,n$, and $(\widehat{\mathbf{F}}_2\widehat{\mathbf{F}}_4^{-1}\widehat{\mathbf{F}}_3)_{ij} = 0$ otherwise. It follows that for $i=1,\ldots,m$

$$(\widehat{\mathbf{T}})_{ij} = \begin{cases} \mathrm{e}^{y_i\delta_j} - \sum_{k=0}^{b-m-1} \frac{(y_i\delta_j)^k}{k!} & j=1,\ldots,n \\ y_i^{b-j} & j=n+1,\ldots,m\,. \end{cases}$$

Note that, for $i=1,\ldots,m$ and $j=1,\ldots,n$,

$$\begin{aligned} (\widehat{\mathbf{T}})_{ij} &= \mathrm{e}^{y_i\delta_j} - \sum_{k=0}^{b-m-1} \frac{(y_i\delta_j)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(y_i\delta_j)^k}{k!} - \sum_{k=0}^{b-m-1} \frac{(y_i\delta_j)^k}{k!} \\ &= \sum_{k=b-m}^{\infty} \frac{(y_i\delta_j)^k}{k!} \\ &= \sum_{h=0}^{\infty} \frac{(y_i\delta_j)^{b-m+h}}{(b-m+h)!} \\ &= (y_i\delta_j)^{b-m} \sum_{h=0}^{\infty} \frac{(y_i\delta_j)^h h!}{(b-m+h)!h!} \\ &= \frac{(y_i\delta_j)^{b-m}}{(b-m)!} \sum_{h=0}^{\infty} \frac{(1)_h(y_i\delta_j)^h}{(b-m+1)_h h!} \\ &= \frac{(y_i\delta_j)^{b-m}}{(b-m)!} {}_1F_1(1; b-m+1; y_i\delta_j) \end{aligned}$$

since $h! = (1)_h$ and $(b-m+h)! = (b-m+1)_h(b-m)!$. Also, for $i=1,\ldots,m$ and $j=n+1,\ldots,m$, $(\widehat{\mathbf{T}})_{ij} = y_i^{b-j} = y_i^{b-m}y_i^{m-j}$.

As a consequence, the matrix $\widehat{\mathbf{T}}$ can be rewritten as $\widehat{\mathbf{T}} = \mathbf{L}\mathbf{G}\mathbf{R}$, where $\mathbf{L}$ and $\mathbf{R}$ are diagonal $m\times m$ matrices given by, respectively, $\mathbf{L} = \mathrm{diag}(y_1^{b-m},\ldots,y_m^{b-m})$, and

$$\mathbf{R} = \mathrm{diag}(\delta_1^{b-m}/(b-m)!,\ldots,\delta_n^{b-m}/(b-m)!,1,\ldots,1)\,.$$

Furthermore, $\mathbf{G}$ is an $m\times m$ matrix whose elements, for $i=1,\ldots,m$, are given by

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-m+1; y_i\delta_j) & j=1,\ldots,n \\ y_i^{m-j} & j=n+1,\ldots,m\,. \end{cases}$$

Thus, we have:

$$\begin{aligned} |\widehat{\mathbf{F}}| &= |\widehat{\mathbf{F}}_4||\widehat{\mathbf{T}}| \\ &= |\mathbf{L}||\mathbf{G}||\mathbf{R}| \prod_{i=0}^{b-m-1} i! \\ &= \frac{|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{b-m}|\mathbf{G}||\mathbf{\Delta}|^{b-m}}{(b-m)!^n} \prod_{i=0}^{b-m-1} i!\,. \end{aligned} \quad (77)$$

In conclusion, by substituting (77) in (76), we get (47).

For $m \le n$, a similar procedure can be used to compute the determinant $|\widehat{\mathbf{F}}|$. In this case,

$$|\widehat{\mathbf{F}}| = \frac{|\mathbf{Y}\mathbf{Y}^{\mathsf{H}}|^{b-n}|\mathbf{G}||\mathbf{\Delta}|^{b-n}}{(b-n)!^m} \prod_{i=0}^{b-n-1} i!\,.$$

Again, by substituting the above expression in (76), we get (46). Here, however, the expression of $(\mathbf{G})_{ij}$ changes as follows:

$$(\mathbf{G})_{ij} = \begin{cases} {}_1F_1(1; b-n+1; y_i\delta_j) & i=1,\ldots,m \\ \delta_j^{n-i} & i=m+1,\ldots,n \end{cases}$$

and for $j=1,\ldots,n$.

APPENDIX I
PROOF OF PROPOSITION 7

The law of the output of a channel, as the one in (23), conditioned on the input power allocation $\mathbf{D}$, is reported in [5, eq. (58)], i.e.,

$$p(\mathbf{Y}|\mathbf{D}) = \frac{\Gamma_n(b)K(\mathbf{Y})|\mathbf{I}+c\gamma\mathbf{D}|^{b-m-1}|\mathbf{F}|}{\pi_n\gamma_n|\gamma c\mathbf{D}|^{b-n}\mathcal{V}(c\mathbf{D})}\,, \qquad (78)$$

where $K(\mathbf{Y}) = \mathrm{e}^{-\|\mathbf{Y}\|^2}/(\pi^{mb}\mathcal{V}(\mathbf{Y}\mathbf{Y}^{\mathsf{H}}))$ and $(\mathbf{F})_{ij} = \exp\left(\frac{c\gamma y_i d_j}{1+c\gamma d_j}\right)$, $i=1,\ldots,n, j=1,\ldots,b$, and $\mathbf{F}_{ij} = y_i^{b-j}$, $i=n+1,\ldots,b, j=1,\ldots,b$.

In order to take average of (78), we first write $|\mathbf{F}|$ as the product of two determinants. Indeed, we partition $\mathbf{F}$ as

$$\mathbf{F} = \left[\begin{array}{cc} \mathbf{F}_1 & \mathbf{F}_2 \\ \mathbf{F}_3 & \mathbf{F}_4 \end{array}\right] \qquad (79)$$

where $(\mathbf{F}_4)_{ij} = y_{n+i}^{b-n-j}\, i,j = 1,\ldots,b-n$, and $\mathbf{F}_1$ is the principle $n\times n$ submatrix of $\mathbf{F}$. Applying the property of the determinant of block matrices [30] to (79), we obtain

$$|\mathbf{F}| = |\mathbf{F}_4||\mathbf{T}|\,, \qquad (80)$$

where $\mathbf{T} = \mathbf{F}_1 - \mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3$. We notice that $|\mathbf{F}_4|$ is independent of $\mathbf{D}$, and the matrix $\mathbf{T}$ has the same size as $\mathbf{D}$. For $m>n$, $p(\mathbf{D})$ is given by (18). We then get

$$\begin{aligned} p(\mathbf{Y}) &= \int p(\mathbf{Y}|\mathbf{D})p(\mathbf{D})\,\mathrm{d}\mathbf{D} \\ &= \frac{\Gamma_n(b)K(\mathbf{Y})|\mathbf{F}_4|}{\pi_n\gamma_n}\int \frac{|\mathbf{I}+c\gamma\mathbf{D}|^{b-m-1}|\mathbf{T}|p(\mathbf{D})}{|\gamma c\mathbf{D}|^{b-n}\mathcal{V}(c\mathbf{D})}\,\mathrm{d}\mathbf{D} \\ &= \frac{\pi_n\Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)}K(\mathbf{Y})|\mathbf{F}_4|}{\gamma_n\Gamma_n(n)c^{n(n-1)/2}\Gamma_n(b-n)\Gamma_n(n+m-b)} \\ &\quad\cdot\int\frac{|\mathbf{I}+c\gamma\mathbf{D}|^{b-m-1}|\mathbf{T}|}{|\mathbf{I}-\mathbf{D}|^{b-m}|\mathbf{D}|^n}\mathcal{V}(\mathbf{D})\,\mathrm{d}\mathbf{D}\,, \\ &= \frac{\pi_n\Gamma_n(b)\Gamma_n(m)(\gamma c)^{n(n-b)}K(\mathbf{Y})|\mathbf{F}_4||\mathbf{Z}|}{\gamma_n c^{n(n-1)/2}\Gamma_n(n)\Gamma_n(b-n)\Gamma_n(n+m-b)} \qquad (81) \end{aligned}$$

where

$$\begin{aligned}(\mathbf{Z})_{ij} &= \int_0^1 \frac{(1+c\gamma x)^{b-m-1}x^{i-1-n}}{(1-x)^{b-m}} \\ &\quad\cdot\left[\exp\left(\frac{c\gamma y_i x}{1+c\gamma x}\right) - (\mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3)_{ij}\right]\end{aligned}$$

has been obtained by using the result in Appendix K, and

$$(\mathbf{F}_2\mathbf{F}_4^{-1}\mathbf{F}_3)_{ij} = \sum_{\ell,k=1}^{b-n}(\mathbf{F}_4^{-1})_{\ell k}\exp\left(\frac{c\gamma x y_{\ell+n}}{1+c\gamma x}\right)y_j^{b-k-n}\,.$$

APPENDIX J
PROOF OF PROPOSITION 8

Conditioned on the input $\mathbf{X}$, the output $\mathbf{Y}$ is complex Gaussian and has i.i.d. rows, so that the evaluation of the differential entropy can be carried out by considering just an arbitrary row, $\mathbf{y}$, of $\mathbf{Y}$ and, then, scaling the result by m. We note that $\mathbf{y}$ is multivariate Gaussian distributed with covariance equal to $(\mathbf{I}+\gamma\mathbf{X}^{\mathsf{H}}\mathbf{X})$. Thus, considering the optimal input matrix, $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\mathbf{\Phi}$ and conditioning on it, the differential entropy is given by:

$$h(\mathbf{y}|\mathbf{X}) = b\log_2(\pi e) + n\mathbb{E}\left[\log_2\left(1+c\gamma\delta\right)\right]\,, \qquad (82)$$

with δ being distributed as a single unordered eigenvalue of the matrix $\mathbf{D}$. By using (18) and considering $p=b-n$ and $q=m+n-b$, $p(\mathbf{D})$ reads as

$$p(\mathbf{D}) = \frac{\pi_n^2\Gamma_n(m)|\mathbf{I}-\mathbf{D}|^{m-b}|\mathbf{D}|^{b-2n}\mathcal{V}^2(\mathbf{D})}{\Gamma_n(n)\Gamma_n(b-n)\Gamma_n(m+n-b)}\,. \qquad (83)$$

By exploiting the result given in Proposition 1 and by denoting the constant terms in the above expression by K, we obtain:

$$p(\delta) = \frac{K}{n}\sum_{i,j=1}^{n}\delta^{(b-2n+i+j-2)}(1-\delta)^{(m-b)}a_{ij} \qquad (84)$$

with a_{ij} being defined as in the above proposition. The integral in (82) can be solved by resorting to partial integration. Indeed, taking $\log_2(1+c\gamma\delta)$ as the primitive factor and recalling that $(1-\delta)^{n-m} = \sum_{\ell=0}^{n-m}\binom{n-m}{\ell}(-1)^\ell\delta^\ell$, by virtue of [31, 3.194.1], we obtain

$$\begin{aligned}\mathbb{E}\left[\log_2\left(1+\gamma\delta\right)\right] &= \frac{K}{n}\sum_{i,j=1}^{n}a_{ij}\sum_{\ell=0}^{m-b}\binom{m-b}{\ell}\frac{(-1)^\ell}{s_{i,j,\ell}-1} \\ &\cdot\left[\log_2(1+c\gamma) - c\gamma\frac{{}_2F_1\left(1,s_{i,j,\ell};s_{i,j,\ell}+1;-\gamma\right)}{\ln(2)(s_{i,j,\ell})}\right]\end{aligned}$$

where $s_{i,j,\ell} = b-2n+i+j+\ell$. Then, using this expression in (82), we get (52).

APPENDIX K
LEMMA 2 IN [32]

Consider a function $\xi(x)$, an arbitrary $n\times n$ matrix $\mathbf{\Phi}(\mathbf{x})$ such that $(\mathbf{\Phi})_{ij} = \phi_i(x_j)$, and an arbitrary $m\times m$ matrix $\mathbf{\Psi}$, $m\geq n$, whose elements are given by

$$(\mathbf{\Psi})_{ij} = \left\{\begin{array}{ll}\psi_i(x_j) & 1\leq i\leq m, 1\leq j\leq n \\ c_{ij} & 1\leq i\leq m, n+1\leq j\leq m\end{array}\right.$$

where c_{ij} are constant. Then, the following identity holds:

$$\int_{[a,b]^n}|\mathbf{\Phi}(\mathbf{x})||\mathbf{\Psi}(\mathbf{x})|\prod_{k=1}^{n}\xi(x_k)\,\mathrm{d}\mathbf{x} = n!|\mathbf{\Xi}| \qquad (85)$$

where, for $1\leq i\leq m$,

$$(\mathbf{\Xi})_{ij} = \left\{\begin{array}{ll}\int_a^b\psi_i(x)\phi_j(x)\xi(x)\,\mathrm{d}x & 1\leq j\leq n \\ c_{ij} & n+1\leq j\leq m\,.\end{array}\right.$$

For the specific case $m=n$, this result appears in [28, Corollary II].

REFERENCES

[1] G. Alfano, A. Lozano, A. Tulino and S. Verdú, "Mutual Information and Eigenvalue Distribution of MIMO Ricean Channels", *International Symposium on Information Theory and its Applications (ISITA),*" Parma, Italy, October 1013, 2004.

[2] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Trans. on Information Theory,* Vol. 56, No. 5, pp. 2307–2359, 2010.

[3] F. Rusek, A. Lozano, and N. Jindal "Mutual Information of IID Complex Gaussian Signals on Block Rayleigh-Faded Channels," *IEEE Trans. on Information Theory,* Vol. 58, No. 1, pp. 331–340, 2012.

[4] B. Hassibi, and T. L. Marzetta, "Multiple-antennas and Isotropically Random Unitary Inputs: the Received Signal Density in Closed Form," *IEEE Trans. on Information Theory*, Vol. 48, No. 6, pp. 1473–1484, 2002.

[5] W. Yang, G. Durisi, and E. Riegler, "On the Capacity of Large-MIMO Block-fading Channels," *IEEE J. Selected Areas in Communications*, Vol. 31, No. 2, pp. 117–132, 2013.

[6] L. Zheng, and D. Tse, "Communicating on the Grassmann Manifold: A Geometric Approach to the Non-coherent Multiple Antenna Channel," *IEEE Trans. on Information Theory*, Vol. 48, No. 2, pp. 359–383, 2002.

[7] S. Moser, "The FadingNumber of IID MIMO Gaussian Fading Channels with a Scalar Line-of-Sight Component," *Annual Allerton Conf. on Comm., Control, and Computing*, Allerton House, Monticello, IL, USA, 2007.

[8] F. Bohagen, P. Orten, and G. E. Oien, "Design of Optimal High-rank Line-of-sight MIMO Channels," *IEEE Trans. on Wireless Communications*, Vol. 6, No. 4, pp. 1420-1425, 2007.

[9] G. Alfano, A. De Maio, and A. M. Tulino, "A Theoretical Framework for LMS MIMO Communication Systems Performance Analysis," *IEEE Trans. on Information Theory*, Vol. 56, No. 11, pp. 5614–5630, 2010.

[10] J. S. Kwak, J. Andrews, A. Lozano, "MIMO Capacity in Correlated Interference-Limited Channels," *IEEE International Symposium on Information Theory (ISIT07)*, Nice, France, June 2007.

[11] A. T. James, "Distribution of Matrix Variates and Latent Roots Derived from Normal Samples," *The Annals of Mathematical Statistics*, Vol. 35, No. 2, pp. 474–501, June 1964.

[12] M.R. McKay, "Random Matrix Theory Analysis of Multiple Antenna Communication Systems," Ph.D. dissertation, Oct. 2006.

[13] A. Ghaderipoor, C. Tellambura, and A. Paulraj, "On the Application of Character Expansions for MIMO Capacity Analysis," *IEEE Trans. on Information Theory*, Vol. 58, No. 5, pp. 2950–2962, 2012.

[14] M. Abramowitz, and I. A. Stegun, "Handbook of Mathematical Functions", *New York: Dover Publications*, 1972.

[15] A. Tulino, and S. Verdú, "Random Matrices and Wireless Communications," *Foundations and Trends in Communications and Information Theory*, Vol. 1, No. 1, July 2004.

[16] A. Gupta, D. Nagar, and A. Vélez-Carvajal, "Unitary invariant and residual independent matrix distributions,", *Computational and Applied Mathematics*, Vol. 28, No. 1, pp. 63–86, 2009.

[17] M. Gursoy, H. V. Poor, and S. Verdú, "Spectral Efficiency of Peak Power Limited Rician Block-Fading Channels," *IEEE International Symposium on Information Theory*, Chicago, IL, June 2004.

[18] G. Taricco, and G. Coluccia, "Optimum Receiver Design for Correlated Rician Fading MIMO Channels with Pilot-aided Detection," *IEEE J. on Selected Areas in Communications*, Vol. 25, No. 7, pp. 1311-1321, Sep. 2007.

[19] K. Liolis, J. Gmez-Vilardeb, E. Casini, and A. I. Prez-Neira, "Statistical Modeling of Dual-Polarized MIMO Land Mobile Satellite Channels," *IEEE Trans. on Communications* Vol. 58, No. 11, pp. 3077–3083, 2010.

[20] D. Hösli and A. Lapidoth, "The Capacity of a MIMO Ricean Channel Is Monotonic in the Singular Values of the Mean," *ITG Conf. Source and Channel Coding*, 2004.

[21] A. K. Nagar and L. Cardeno, "Matrix-variate Kummer-Gamma distribution," *Random Operators and Stochastic Equations*, Vol. 9, No. 3, pp. 208–217, 2001.

[22] T. L. Marzetta, and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Information Theory*, Vol. 45, No. 1, pp. 139–157, 1999.

[23] A. Lapidoth, and S. M. Moser, "Capacity Bounds via Duality with Applications to Multiple-Antenna Systems on Flat Fading Channels," *IEEE Trans. on Inf. Theory*, Vol. 49, No. 10, pp. 2426–2467, 2003.

[24] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. on Wireless Communications*, Vol. 9, No. 11, pp. 3590–3600, 2010.

[25] C. G. Khatri, "On Certain Distribution Problems Based on Positive Definite Quadratic Functions in Normal Vectors," *The Annals of Mathematical Statistics*, Vol. 37, No. 2, pp. 467–479, Apr. 1966.

[26] M. McKay and I. Collings, "General Capacity Bounds for Spatially Correlated Rician MIMO Channels", *IEEE Transactions on Information Theory*, Vol. 51, No. 9, September 2005.

[27] G. Alfano, A. Tulino, A. Lozano, and S. Verdú, "Eigenvalue Statistics of Finite-Dimensional Random Matrices for MIMO Wireless Communications," *IEEE ICC*, Istanbul, Turkey, June 2006.

[28] M. Chiani, M. Z. Win, and A. Zanella, "On the Capacity of Spatially Correlated MIMO Rayleigh-Fading Channels," *IEEE Trans. on Information Theory*, Vol. 49, No. 10, pp. 2363–2371, 2003.

[29] C. Itzykson, and J. B. Zuber, "The Planar Approximation. II," *J. of Mathematical Physics*, Vol. 21, No. 3, pp. 411-421, 1980.

[30] R. A. Horn, and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[31] I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.

[32] H. Shin, M. Win, J. Lee, and M. Chiani, "On the Capacity of Doubly Correlated MIMO Channels," *IEEE Transactions on Wireless Communications*, Vol. 5, No. 8, pp. 2253–2265, August 2006.

Giusi Alfano received the Laurea degree in communication engineering from the University Federico II, Naples, Italy, in 2001, where she was CNIT junior researcher from 2003 to 2005, and the phd in Information Engineering from University of Sannio, Benevento, Italy, in 2007. She was visiting researcher at ftw, Wien, in 2007 and at Chalmers University, Goteborg, in 2012. She was post-doc in 2009 in the Alcatel-Lucent Chair of Flexible Communications, Supélec, Paris, and in 2011 she was ERCIM post-doc in NTNU, Trondheim, Norway. She currently holds a post-doctoral position at Politecnico di Torino, Italy. Her research work lies mainly in the field of random matrix theory applications to MIMO wireless communications and sensor networks.

Carla-Fabiana Chiasserini (M'98, SM'09) graduated in Electrical Engineering (summa cum laude) from the University of Florence in 1996. She received her Ph.D. from Politecnico di Torino, Italy, in 2000. She has worked as a visiting researcher at UCSD in 1998–2003, and she is currently an Associate Professor with the Department of Electronic Engineering and Telecommunications at Politecnico di Torino. Her research interests include architectures, protocols, and performance analysis of wireless networks. Dr. Chiasserini has published over 200 papers in prestigious journals and leading international conferences, and she serves as Associated Editor of several journals.

Alessandro Nordio (S'00-M'03) is currently a researcher with the Institute of Electronics, Computer and Telecommunication Engineering of the Italian National Research Council. In 2002 he received the Ph.D. in Telecommunications from "Ecole Polytechnique Federale de Lausanne", Lausanne, Switzerland. From 1999 to 2002, he performed active research with the Department of Mobile Communications at Eurecom Institute, Sophia Antipolis (France). From 2002 to 2009 he was a post-doc researcher with the Electronic Department of Politecnico di Torino, Italy. His research interests are in the field of signal processing, space-time coding, wireless sensor networks and theory of random matrices.

Siyuan Zhou (S'14) received the B.S. degree in electronic engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2008, the M.S. degree in communication and information systems from Southeast University, Nanjing, China, in 2011. Since Jan. 2012, he has been pursuing his Ph.D. degree at the Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy. His research lies mainly in the field of random matrix theory applications to MIMO wireless communication systems.